QUANTITATIVE ERGODIC THEOREMS AND THEIR NUMBER-THEORETIC APPLICATIONS

ALEXANDER GORODNIK AND AMOS NEVO

Abstract. We present a survey of ergodic theorems for actions of algebraic and arithmetic groups recently established by the authors, as well as some of their applications. Our approach is based on spectral methods employing the unitary representation theory of the groups involved. This allows the derivation of ergodic theorems with a rate of convergence, an important phenomenon which does not arise in classical ergodic theory. Quantitative ergodic theorems give rise to new and previously inaccessible applications, and we demonstrate the remarkable diversity of such applications by presenting several number-theoretic results. These include, in particular, general uniform error estimates in lattice points counting problems, explicit estimates in sifting problems for almost-prime points on symmetric varieties, bounds for exponents of intrinsic Diophantine approximation, and results on fast distribution of dense orbits on homogeneous spaces.

Contents

1. Introduction

The goals of the present survey are to describe some mean and pointwise ergodic theorems for actions of algebraic groups and their lattice subgroups, and to give an exposition of some of the diverse applications that such ergodic theorems have, particularly to counting, equidistribution and Diophantine approximation problems in number theory. The results described below have been recently developed by the authors, and constitute an expansion of the range of applications of spectral methods in ergodic theory, particularly in the context of dynamics on homogeneous space. The survey is however limited in scope in that it describes only the authors'

Date: April 26, 2013.

specific spectral approach, and it does not discuss other approaches within the very large and rapidly expanding field of homogeneous dynamics and its application in number theory.

In order to situate our approach in the proper context, we start by describing below some elements in the interesting story of the development of the mean ergodic theorem, beginning with its formulation by von-Neumann in 1932. Let us however note here very briefly the main new points in our discussion. First, the ergodic theorems, which we establish, are developed using spectral methods and representation theory of non-amenable algebraic groups, subjects left out of most of traditional ergodic theory. Second, these ergodic theorems provide a rate of convergence to the ergodic limit, a remarkable and most useful phenomenon that does not arise in classical ergodic theory. Third, the rate of convergence allows new and previously inaccessible applications of ergodic theorems to be developed, using the fact that the actions of the groups we study are closely connected to a variety of very natural number-theoretic problems.

To demonstrate the utility and diversity of such applications, let us briefly describe a selection of counting, equidistribution, and Diophantine approximation problems that may at first sight appear completely unrelated. The main rationale of our exposition in the present survey is the fact, which will be explained in detail below, that solutions to all of the following problems depend on applying a suitable ergodic theorem to the case under consideration.

1. The uniform lattice point counting problem. Let Γ be a discrete subgroup of a locally compact group G such that the space G/Γ carries a finite G-invariant measure. The lattice point counting problem is to count the number of elements of Γ in an increasing family of bounded domains $B_t \subset G$, and ideally to establish an asymptotic formula

$$|\Gamma \cap B_t| = \frac{|B_t|}{|G/\Gamma|} + O(|B_t|^\kappa)$$

with $\kappa < 1$ as small as possible. Some questions that arise here are, for example, what is the asymptotics of the number of unimodular integral matrices A lying in the norm ball $\|A\| < t$ as $t \to \infty$, or in more general domains in the matrix space ? What is the asymptotics for the number of such matrices A satisfying a congruence condition $A = A_0 \mod q$? More generally, what is the asymptotics of the number of unimodular matrices with entries in the ring of integers of an algebraic number field, which are of bounded height?

Counting lattice points is in fact the most basic of all mean ergodic theorems that we consider. This point of view that will be explained in full in §5 below, where we will describe a general solution to the lattice point counting problem.

2. Almost prime points on algebraic varieties. As we shall see below, establishing robustness and uniformity properties of the solution to the lattice point counting problem is crucial in several important applications, which depend on answering more refined counting questions. Of those, let us mention the problem of existence of *almost prime* integral points on homogeneous varieties, raised in [NS10]. Consider for example the set of symmetric integral positive-definite matrices of fixed determinant, and call such a matrix prime if all its entries are prime

numbers. How large is the set of prime points in the set of integral points? Less ambitiously, how large is the set of integral points all of whose entries are r-almost primes (i.e., products of at most r prime factors) ?

This question has been the subject of intensive activity in recent years, and we refer to [BGS10] and the recent surveys [Lu12] and [K] for further discussion. In §6 below we will utilise uniformity in the solution of the lattice point counting problem to establish an asymptotic lower bound of the right magnitude for the number of r-prime points lying on symmetric varieties.

3. Diophantine approximation on algebraic varieties. Consider the rational ellipsoid $\{(x,y,z)\,;\, ax^2+by^2+cz^2=d\}$, $a,b,c,d\in\mathbb{N}_+$. Assuming that the rational points on the ellipsoid are dense, is it possible to establish a rate of approximation of a general points on the ellipsoid by rational points ? Is it possible to establish a rate of approximation by rational points satisfying additional integrality constraints ?

These problems, in the much more general context of homogeneous algebraic varieties, were raised in S. Lang's 1965 Report on Diophantine Approximation [L65]. We will state this problem precisely in §7 and describe a general solution, obtained jointly with A. Ghosh, in the case where the variety is homogeneous under an action a simple algebraic group. For the rational ellipsoids we establish a rate of approximation by rational points satisfying integrality constrains that is best possible in a natural sense.

4. Equidistribution of typical orbits on homogeneous spaces. Consider the one-sheeted hyperboloid $\{(x,y,z)\,;\, ax^2+by^2-cz^2=d\}$, $a,b,c,d\in\mathbb{N}_+$, which is also known as de-Sitter space. The group of integral matrices preserving the corresponding quadratic form naturally acts on this space, and typical orbits are dense, but does a typical orbit have a limiting distribution in a suitable sense ?

This problem, for general lattices in the isometry group of a Lorentz form acting on de-Sitter space of arbitrary dimension, was raised by V. Arnold in 1985 [Ar]. A precise quantitative solution for general homogeneous varieties will be described in §9.

5. Equidistribution in algebraic number fields. The group $\Gamma=\mathrm{SL}_2(\mathbb{Z}[\sqrt{d}])$, where $d>0$ is not a square, is a dense subgroup of $G=\mathrm{SL}_2(\mathbb{R})$. Consider elements of Γ of bounded norm : $\Gamma_t=\{\gamma\in\Gamma:\ \log(\|\gamma\|^2+\|\bar{\gamma}\|^2)\le t\}$, where $\bar{\gamma}$ denotes the Galois conjugate. Do these sets become equidistributed in G in a suitable sense ?

We note that the action of Γ on G by translations is isometric, and the natural problem is to prove equidistribution for *every single orbit*, i.e. every coset of Γ. One possible such statement is a ratio equidistribution theorem, establishing that for nice domains $\Omega_1,\Omega_2\subset G$,

$$\frac{|\Gamma_t\cap\Omega_1|}{|\Gamma_t\cap\Omega_2|}\to\frac{|\Omega_1|}{|\Omega_2|}\quad \text{as } t\to\infty\,,$$

with the convergence taking place at a definite rate. Such a result will be stated precisely and explained in §9.

We remark that the problem of establishing a ratio equidistribution theorem for dense subgroups acting by isometries on the group or its homogeneous spaces

was raised originally by Kazhdan [K67], in the case of a dense group of Euclidean isometries acting on the Euclidean plane. We refer to [V] for a recent progress on this problem.

6. Ergodic theorems for lattice subgroups. Finally, let us mention some natural questions which constitute direct generalizations of the fundamental convergence problems of classical ergodic theory. Taking the group $\Gamma = \mathrm{SL}_n(\mathbb{Z})$ as an example, consider its linear action on the torus $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$. Does the mean ergodic theorem hold for the action of Γ on $\mathbb{T}^n$? Specifically, let $\Gamma_t = \{\gamma \in \Gamma\,;\, \log\|\gamma\| \le t\}$, and for $f \in L^p(\mathbb{T}^n)$ and $x \in \mathbb{T}^n$, consider the averages

$$\frac{1}{|\Gamma_t|}\sum_{\gamma\in\Gamma_t} f(\gamma^{-1}x).$$

Do these averages converge to the space average $\int_{\mathbb{T}^n} f(x)\,dx$ in L^p? Does convergence hold almost everywhere ? If so, is there a rate of convergence to the space average, in L^p and almost everywhere ?

These questions also arise for other interesting actions of $\mathrm{SL}_n(\mathbb{Z})$ (for instance, for the action on the space of unimodular lattices or for the action on its profinite completion), and they are all resolved by a general pointwise and mean ergodic theorem for the averages on norm balls in $\mathrm{SL}_n(\mathbb{Z})$, a result that will be formulated more generally and precisely in §8 and explained there.

To conclude the introduction, let us emphasize that it is ultimately the existence of an explicit rate of convergence in the ergodic theorems that is responsible for solution of the above problems. We will therefore give in our discussion high priority to the problem of establishing an explicit rate of convergence of the averages we consider to the space average in measure-preserving dynamical systems.

To situate our approach in its proper context, let us now turn to describe some elements in the development of the mean ergodic theorem. We will also mention some of the milestones in the applications of spectral methods in ergodic theory. As we shall see, an interesting feature of this story is that the success of the classical methods of ergodic theory is inextricably tied to both the impossibility of establishing a rate of convergence in the ergodic theorems, and the impossibility of extending the classical methods to classes of groups such as simple algebraic groups or their lattice subgroups.

2. Averaging in ergodic theory

2.1. **The classical mean ergodic theorem.** The basic objects of study in classical Hamiltonian mechanics consist of a compact Riemannian manifold M, which is called the phase space, together with a divergence-free vector field on M. We denote by μ the Riemannian volume on M, normalised to be a probability measure. The integral curves of the vector field give rise to a one-parameter group of volume-preserving transformations $a_s : M \to M$, $s \in \mathbb{R}$, which describes the time evolution in phase space. For a function $f : M \to \mathbb{R}$, the time averages along orbits

are defined by

$$\mathcal{A}_t f(x) = \frac{1}{t} \int_0^t f(a_s x)\, ds.$$

One of the most fundamental questions in classical dynamics is the elucidation of their long term behavior. Two important questions that arise are :

- do the time averages $\mathcal{A}_t f$ converge at all, and if so, what is their limit?
- how does the limit depend on the initial state x and the function f?

The solution proposed by von Neumann to this problem was to consider for every measure-preserving map $a_s : M \to M$ an associated unitary transformation, the Koopman operator $U_s : L^2(M) \to L^2(M)$, given by $U_s f(x) = f(a_s x)$. Thus the measure-preserving flow gives rise to a one-parameter group of unitary transformations. This made it possible for von-Neumann to apply his recently developed spectral theory to prove the mean ergodic theorem [vN32]:

Theorem 2.1 (von Neumann)**.** *Let (X, μ) be a (standard) probability space with a one-parameter flow of measure-preserving transformations $a_s : X \to X$, $s \in \mathbb{R}$. Then for any $f \in L^2(X)$,*

$$\mathcal{A}_t f \to \mathcal{P} f \quad \text{as } t \to \infty,$$

in L^2-norm, where $\mathcal{P}$ denotes the orthogonal projection on the space of functions invariant under the flow.

In particular, consider the case where the space of invariant functions in $L^2(M)$ consists of constant functions only, or equivalently, every invariant measurable set has measure zero or one. Flows satisfying this property are called *ergodic*. Then the orthogonal projection of f to the space of constant functions is given by the "space average" $\int_X f\, d\mu$. Thus, for ergodic flows the time averages converge in L^2-norm to the space average:

$$\mathcal{A}_t f \to \int_X f\, d\mu \quad \text{as } t \to \infty,$$

verifying the "ergodic hypothesis".

Several years later, F. Riesz's [R38] gave an elegant elementary proof of the mean ergodic theorem, which became very influential in the subsequent development of ergodic theory. We will indicate it here in full, in order to highlight its main idea. Let $U_s : \mathcal{H} \to \mathcal{H}$, $s \in \mathbb{R}$, be a one-parameter group of unitary operators on a Hilbert space $\mathcal{H}$ and $\mathcal{P}$ denote the projection on the space of vectors invariant under the group. We will show that the averaging operators $\mathcal{A}_t = \frac{1}{t}\int_0^t U_s\, ds$ satisfy $\mathcal{A}_t f \to \mathcal{P} f$ for every $f \in \mathcal{H}$ by breaking $\mathcal{H}$ up to two orthogonal complementary subspaces and arguing separately in each of them. First, we apply $\mathcal{A}_t$ to vectors f of the form $f = U_{t_0} h - h$. Then

$$\begin{aligned}
\mathcal{A}_t f = \mathcal{A}_t(U_{t_0} h - h) &= \frac{1}{t}\left(\int_{t_0}^{t+t_0} U_s h\, ds - \int_0^t U_s h\, ds\right) \\
&= \frac{1}{t}\left(\int_t^{t+t_0} - \int_0^{t_0}\right) U_s h\, ds,
\end{aligned}$$

and therefore

$$\|\mathcal{A}_t f\| \leq \frac{2t_0}{t}\,\|h\| \to 0 \quad \text{as } t \to \infty.$$

Since f is orthogonal to every invariant vector, namely $\mathcal{P}f = 0$, the mean ergodic theorem holds for f as stated, hence also for the closed linear span of such f's. Obviously, the mean ergodic theorem holds in the space of invariant vectors. The proof is therefore complete upon observing that the span of $U_{t_0}h - h$ with $t_0 \in \mathbb{R}$ and $h \in \mathcal{H}$ is dense in the orthogonal complement of the space of invariant vectors.

Notice that the main ingredient of the above proof is the property of asymptotic invariance (under translation) of the intervals $[0,t]$, namely the fact that the measure of the symmetric difference of $[0,t]$ and $[0,t]+t_0$ divided by the measure of $[0,t]$ converges to zero at $t\to\infty$. This crucial property of intervals on the real line admits a straightforward generalisation. Let G be a locally compact second countable group G equipped with a left-invariant Haar measure. Define a one-parameter family of sets $F_t \subset G$ of positive finite measure to be *asymptotically invariant* if it satisfies

$$\frac{|F_t g \triangle F_t|}{|F_t|} \to 0 \qquad \text{as } t \to \infty. \tag{2.1}$$

uniformly for g in compact sets of G. A group admitting such an asymptotically invariant family (also called a *Følner family*) is called *amenable*.

Given a measure-preserving action of the group G on a probability space (X,μ), define for $g \in G$ the unitary Koopman operators

$$\pi_X(g) : L^2(X) \to L^2(X) : f \mapsto f(g^{-1}x).$$

It is straightforward to adapt Riesz's argument and conclude that amenable groups satisfy the mean ergodic theorem. Namely, denoting by β_t the uniform probability measures supported on an asymptotically invariant (Følner) family F_t and by $\pi_X(\beta_t)$ the corresponding averaging operators in $L^2(X)$, we conclude that

$$\pi_X(\beta_t)f = \frac{1}{|F_t|}\int_{F_t} \pi_X(g)f\,dg \to \mathcal{P}f \quad \text{as } t \to \infty,$$

in L^2-norm.

2.2. **Equidistribution and pointwise convergence.** So far we have discussed the convergence of the averages $\pi_X(\beta_t)f$ in L^2-norm. The question as to the convergence $\pi_X(\beta_t)f(x)$ for individual points $x \in X$ is of fundamental importance as well.

We begin by describing the best-case scenario, namely equidistribution, and its characterisation in amenable dynamics. Let X be a compact metric space on which G acts continuously, and let $F_t \subset G$ be a family of Følner sets that define the corresponding averaging operators $\pi_X(\beta_t)$. An extreme possibility is that for every continuous function f and for every point $x \in X$, the "time averages" $\pi_X(\beta_t)f(x)$ converge uniformly to a limit $m(f)$, which is independent of x. In that case, the functional $f \mapsto m(f)$ defines a a probability measure m on X, which is invariant under G. Furthermore m is then the unique G-invariant probability measure on X, and the system is called *uniquely ergodic*. Conversely, when there exists a unique

G-invariant probability measure on X, the "time averages" of f converge uniformly to the space average $\int_X f\,dm$.

In general the situation is far more complicated, and the limit of the averages $\pi_X(\beta_t)f(x)$ depends very sensitively on the initial point x. Typically there exist two (or even uncountably many) mutually singular invariant probability measures μ_1 and μ_2 on X, such that the following holds. The time averages $\pi_X(\beta_t)f(x)$ of a continuous function f converge on a set X_1 of full μ_1-measure to $\int_X f d\mu_1$, and on a set X_2 of full μ_2-measure to $\int_X f d\mu_2$, with the sets X_1 and X_2 disjoint, and $\int_X f d\mu_1 \neq \int_X f d\mu_2$. The question of pointwise convergence for one-parameter flows is addressed by Birkhoff's pointwise ergodic theorem [B32]:

Theorem 2.2 (Birkhoff)**.** *Let (X,μ) be a (standard) probability space with a one-parameter flow of measure-preserving transformations $a_s : X \to X$, $s \in \mathbb{R}$. Then for any $f \in L^1(X)$, the time averages $\frac{1}{t}\int_0^t f(a_s x)\,ds$ converge for μ-almost every $x \in X$ to $\mathcal{P}f(x)$.*

An elementary proof of Birkhoff's theorem which relies on asymptotic invariance of intervals as well as their order structure was given by [KW82]. So far, this proof has not been generalized to other groups, including $\mathbb{Z}^2$. The most successful approach to the proof of poinwise ergodic theorems relies on controlling the maximum deviation between the time averages and the space average throughout the entire sampling process. Thus we consider the following fundamental quantity, called the *maximal function* associated to the flow :

$$\mathcal{M}f = \sup_{t>0} \left|\pi_X(\beta_t)f - \mathcal{P}f\right|.$$

In order to establish the pointwise convergence, we seek to prove the following two properties of the flow:

I. Existence of a dense set of functions $f \in L^1(X)$ for which $\pi_X(\beta_t)f(x)$ converges for almost every $x \in X$.

II. The maximum error in the measurement performed by the time averages is controlled by the *strong maximal inequality* in $L^2(X)$:

$$\left\|\mathcal{M}f\right\|_2 \leq \text{const} \cdot \left\|f\right\|_2,$$

or by the *weak maximal inequality* in $L^1(X)$:

$$\mu(\{\mathcal{M}f > r\}) \leq \text{const} \cdot r^{-1} \left\|f\right\|_1.$$

Once these properties are established, it follows by a classical approximation argument known as the Banach principle that the time averages converge almost everywhere for every $f \in L^1(X)$.

The proof of the first property proceeds by using asymptotic invariance of a Følner family, which implies that pointwise convergence holds for the family of functions $\pi_X(g)f - f$ with $f \in L^\infty(X)$ and $g \in G$. This is a straightforward adaptation of Riesz's argument described above.

The proof of the maximal inequality utilizes two separate ideas : the *transference principle* and *covering arguments*. The transference principle asserts that the maximal inequality for any family of operators $\pi_X(\beta_t)$ acting on a general space (X,μ) can be reduced to the maximal inequality for the convolution operators defined by

β_t on the group G. This principle was initiated by Wiener [W39], and generalized by Calderon [C53] and Coifman–Weiss [CW76]. Finally, the investigation of the convolution operators exploits *geometric covering properties* of the translates of sets in a Følner family F_t. This approach originated with Wiener [W39], who has generalised the Hardy-Littlewood maximal inequality from the real line to Euclidean space by introducing an elegant and influential covering argument based on the *doubling property*

$$|B_{2t}| \leq \text{const} \cdot |B_t|$$

for volumes of the Euclidean balls B_t.

The pointwise ergodic theorem was subsequently established, in turn, for groups with polynomial volume growth, exponential solvable Lie groups, and connected amenable Lie groups. This was carried out by Templeman [T67], Emerson–Greenleaf [EG74], Coifman–Weiss [CW76] and others. It should be pointed out, however, that a Følner sequence may fail to satisfy the pointwise ergodic theorem, even for the group $\mathbb{Z}$. The pointwise ergodic theorem for general amenable groups requires a regularity assumption stronger than asymptotic invariance, namely temperedness, and was established by Lindenstrauss in [L01]. We refer to [N06] for a recent survey and to [AAB] for a detailed account of the ergodic theory of amenable groups.

3. From amenable to non-amenable groups

The proofs of the ergodic theorems indicated in the previous section were founded upon the almost invariance property (2.1) of the Følner families. How then can one proceed when the group is not amenable, so that there are no asymptotically invariant families of sets at all?

Before describing our approach, let us note one more consequence of the existence of asymptotically invariant sequences of sets in the group. For any properly ergodic (i.e., not transitive) action of a countable amenable group G, one can show that the space X admits a non-trivial asymptotically invariant sequence of sets, namely, a sequences of measurable sets $A_n \subset X$, such that

$$0 < \liminf_{n\to\infty} \mu(A_n) \leq \limsup_{n\to\infty} \mu(A_n) < 1 \quad \text{and} \quad \mu(gA_n \triangle A_n) \to 0$$

uniformly for g in compact subsets of G. It then follows easily that there exists an *asymptotically invariant sequence* of functions $f_n \in L^2(X)$ with zero integral such that

$$\|f_n\| = 1 \quad \text{and} \quad \|\pi_X(g)f_n - f_n\| \to 0 \tag{3.1}$$

uniformly for g in compact subsets of G. In fact, the existence of an asymptotically invariant sequence in every property ergodic action characterises amenable groups (see [S81, JR79]).

We can therefore conclude that a countable group G is amenable if and only if in every properly ergodic action, the averaging operators satisfy

$$\left\| \pi_X(\beta)|_{L^2_0(X)} \right\| = 1 \tag{3.2}$$

for every probability measure β on G, where $L^2_0(X)$ denotes the orthogonal complement to the space of constant functions. In particular, it follows that for a properly ergodic action of a countable amenable group, *no uniform rate of convergence* in

the mean ergodic theorem can be established, namely the norm of the averaging operators does not decay at all.

3.1. **Spectral gap property.** Our discussion so far has lead us to the following observation : when G is a countable non-amenable group, at least some actions X have the property that the operators $\pi_X(\beta)$ are strict contractions on $L^2_0(X)$, namely $\|\pi_X(\beta)\|_{L^2_0} < 1$, for all absolutely continuous generating probability measure β on G. We can therefore conclude that a non-amenable countable group G necessarily has actions which admit a spectral gap, in the sense of the following general definition.

Definition 3.1. A measure-preserving ergodic action of a group G on a probability space (X,μ) has a *spectral gap* if one of the following two equivalent conditions holds:

- There is no asymptotically invariant sequences of functions as in (3.1).
- There exists an absolutely continuous symmetric probability measure β whose support generates G such that
$$\left\| \pi_X(\beta)|_{L^2_0(X)} \right\| < 1. \tag{3.3}$$

We remark that it follows that for actions with spectral gap the estimate (3.3) holds for every absolutely continuous probability measure β on G whose support generates G as a semigroup.

Kazhdan [K67] has made the truly fundamental discovery that many groups satisfy an extremely general form of the spectral gap condition. Not only it is satisfied for all ergodic probability preserving actions of the group, an analogous property holds in fact for all unitary representations without invariant unit vectors.

Definition 3.2. A group G has *Kazhdan property (T)* if one of the following equivalent conditions holds:

- Every unitary representation π of G which has no invariant unit vectors also has no sequences of asymptotically invariant unit vectors as in (3.1).
- For some absolutely continuous symmetric probability measure β on G whose support generates G,
$$\sup_\pi \|\pi(\beta)\| < 1, \tag{3.4}$$
where the supremum is taken over all unitary representations of G that have no invariant unit vectors.

Examples of groups satisfying property (T) abound. For instance, the groups $\mathrm{SL}_n(\mathbb{R})$, $\mathrm{SL}_n(\mathbb{C})$ and $\mathrm{Sp}_n(\mathbb{R})$ with $n \geq 3$ have this property. Among simple Lie groups, only the isometry groups of real hyperbolic spaces and the isometry groups of complex hyperbolic spaces do not have property (T) (the case of rank-one groups is dealt by [K69]). Similarly, simple algebraic groups over non-Archemedian local fields all have property (T), unless their split rank is one. We refer to the monograph [BHV] for an extensive discussion of this property.

The isometry groups of real and complex hyperbolic spaces (and algebraic group of split rank 1) do in fact have probability preserving actions admitting an asymptotically invariant sequence of unit vectors, namely actions which do not have a

spectral gap. Nevertheless, many of the most important and natural actions of these groups do have a spectral gap, although it may vary from action to action, and it does not obey the remarkable uniform estimate provided by property T. The problem of establishing uniform spectral estimates for certain families of actions (or families of unitary representations) of these groups is a problem of foundational importance, which has been studied extensively. Let us now turn to describe some of the results and the challenges of establishing such spectral estimates uniformly, and then explain the utility of such estimates in proving ergodic theorems with rates of convergence for actions of semisimple algebraic groups.

3.2. **Selberg's property and property τ.** The phenomenon where a family of actions with increasing complexity satisfies the spectral gap property uniformly, namely the estimate in (3.3) is uniform, is termed *property* (τ).

The first instance of the property (τ) was discovered by A. Selberg in [S65] for congruence subgroups of $\mathrm{SL}_2(\mathbb{Z}) \subset \mathrm{SL}_2(\mathbb{R})$. Let us consider a sequence of covers of finite area hyperbolic surfaces

$$\mathbb{H}^2 \to \cdots \to S_n \to \cdots \to S_2 \to S_1 = \mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}^2.$$

We denote by $\lambda(S_n)$ the bottom of the spectrum (excluding the zero eigenvalue) of the Laplace operator $\Delta = y^2\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)$ on S_n. One might expect that when the covers S_n grow in a regular fashion, they "approximate" the hyperbolic space $\mathbb{H}^2$, and $\lambda(S_n)$ approaches $\lambda(\mathbb{H}^2) = \frac{1}{4}$. In particular, $\lambda(S_n)$ is expected to stay bounded away from zero. Here the assumption that covers grow in a regular fashion is essential, and A. Selberg constructed an example of a sequence of covers S_n such that $\lambda(S_n) \to 0$ as $n \to \infty$. On the other hand, for the surfaces $S_\ell = \Gamma(\ell)\backslash\mathbb{H}^2$, where $\Gamma(\ell) = \{\gamma \in \mathrm{SL}_2(\mathbb{Z}) : \gamma = \mathrm{I} \mod \ell\}$ denotes the congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$ of level ℓ, he showed that

$$\lambda(S_\ell) \geq \frac{3}{16}. \tag{3.5}$$

Moreover, A. Selberg conjectured that for the congruence covers,

$$\lambda(S_\ell) \geq \lambda(\mathbb{H}^2) = \frac{1}{4}. \tag{3.6}$$

Although estimate (3.5) has been improved in [GJ78, LRS95, I96, KS03], the Selberg conjecture (3.6) is still open. We refer to the surveys [S95, S03] for a more detailed discussion. Nowadays this conjecture is understood as a special case of the generalised Ramanujan conjectures, discussed in [S05, BB13], which are most conveniently stated in the representation-theoretic terms. These conjectures also include the original Ramanujan conjecture concerning the bounds on Fourier coefficients of holomorphic modular forms, which was solved by P. Deligne.

More generally, let $\mathrm{G} \subset \mathrm{GL}_d(\mathbb{C})$ be a semisimple algebraic group defined over $\mathbb{Q}$, $\Gamma = \mathrm{G}(\mathbb{Z})$, and let

$$\Gamma(\ell) = \{\gamma \in \Gamma : \gamma = \mathrm{I} \mod \ell\}$$

denote the congruence subgroups. We denote by

$$\mathcal{H}_{\mathrm{G}}(\ell) = L^2_0(\mathrm{G}(\mathbb{R})/\Gamma(\ell))$$

the space of square-integrable functions with zero integral on the space $\mathrm{G}(\mathbb{R})/\Gamma(\ell)$. We consider the family of unitary representation of $\mathrm{G}(\mathbb{R})$ on $\mathcal{H}_{\mathrm{G}}(\ell)$ defined by

$$\pi_\ell(g)\phi(x) = \phi(g^{-1}x), \quad \phi \in \mathcal{H}_{\mathrm{G}}(\ell). \tag{3.7}$$

The generalised Ramanujan conjectures seeks to describe the decomposition of the representations π_ℓ (and more general automorphic representations) into irreducible constituents. We refer to [S05, BB13] for a comprehensive survey of this subject, but what would be of crucial importance for our purposes is the uniform spectral gap property generalizing (3.5). In the representation-theoretic language, this property can be reformulated in terms of isolation of the trivial representation of $\mathrm{G}(\mathbb{R})$ from the other irreducible representations appearing in π_ℓ for all $\ell \geq 1$ or in terms of estimates on the integrability exponents of the representations π_ℓ, which we proceed to define.

If Π is a family of unitary representations, we say that Π has *property* (τ) if (3.4) holds with $\pi \in \Pi$. It is of crucial for many number-theoretic applications to know the uniform spectral gap property for the family of automorphic representations (3.7). This problem has a long history and is closely connected with works on Langlands functoriality conjectures. In particular, for other forms of the group SL_2, property (τ) follows from Gelbart–Jacquet correspondence [GJ78]. An important milestone was reached by M. Burger and P. Sarnak [BS91] who showed that it is sufficient to verify property (τ) for certain proper subgroups of the ambient group. Combined with previous work, this established property (τ) for all semisimple $\mathbb{Q}$-simple simply connected algebraic groups except the unitary groups. Finally, the case of unitary groups was settled by L. Clozel in [C03].

In the case of semisimple Lie groups, it is convenient to measure the quality of spectral gap in terms of integrability exponents that were introduced by Cowling [C79], and Howe and Moore [HM79], as follows. Let π be a unitary representation of a locally compact group G on a Hilbert space $\mathcal{H}$. For $v, w \in \mathcal{H}$, we denote by $c_{v,w}(g) = \langle \pi(g)v, w\rangle$ the corresponding matrix coefficient, which is a continuous bounded function on G. We define the *integrability exponent* of the representation π by

$$q(\pi) = \inf\{q > 0 : \, c_{v,w} \in L^q(G) \text{ for all } v, w \text{ in a dense subset of } \mathcal{H}\}. \tag{3.8}$$

It was shown in [BW80, C79] that when G is a connected simple Lie group with property T,

$$\sup_\pi q(\pi) < \infty, \tag{3.9}$$

where the supremum is taken over all unitary representations of G that have no nonzero invariant vectors. The value of the supremum has been estimated in [HT, L95, LZ96, O02, N03]. The estimate (3.9) can be considered as a quantitative version of property (T). We also note the important fact that the works on property (τ) for automorphic representations also lead to explicit estimates on $\sup_{\ell\geq 1} q(\pi_\ell)$.

4. Spectral gaps and ergodic theorems

4.1. **Quantitative mean and pointwise theorems for G-actions.** We now turn to discuss the applications of the spectral estimates described in the previous section in the proofs of ergodic theorems. Let (X, μ) be a (standard) probability

space with a measure-preserving action of a locally compact second countable G equipped with left-invariant Haar measure. For a family of measurable sets B_t of G with finite positive measure, we denote by β_t the uniform probability measures supported on B_t. We aim to prove ergodic theorems for the corresponding averaging operators $\pi_X(\beta_t)$ defined by

$$\pi_X(\beta_t)f(x) = \frac{1}{|B_t|}\int_{B_t} f(g^{-1}x)\,dg, \quad f \in L^p(X). \tag{4.1}$$

In the rest of this section, we discuss quantitative ergodic theorems for connected simple Lie groups G (for instance, $G = \mathrm{SL}_n(\mathbb{R})$). In order to prove the quantitative mean ergodic theorem for the averages $\pi_X(\beta_t)$, we need to estimate the norms of the operators $\pi_X(\beta_t)|_{L^2_0(X)}$, and this is achieved using the method of "spectral transfer", originating in [N98]. The methods consists of two separate steps. First, it is possible to reduce the norm estimate of $\pi(\beta_t)$ in a general representation π, to a norm estimate for the convolution operator $\lambda_G(\beta_t)$ for the regular representation λ_G of G acting on $L^2(G)$. This serves as an analogue in this context, of the transference principle for amenable groups. Second, it is then necessary to estimate the norm of the convolution operators on G.

Put in quantitative form, the principle of spectral transfer is based on the following two fundamental facts:

- If the action of G on (X,μ) has spectral gap, then the representation $\pi_X|_{L^2_0(X)}$ has finite integrability exponent, and it follows that sufficiently high tensor power representation $(\pi_X|_{L^2_0(X)})^{\otimes n}$ embeds as a subrepresentation of the regular representation λ_G of G. It then follows from Jensen's inequality that for even n,

$$\left\| \pi_X(\beta_t)|_{L^2_0(X)} \right\| \le \|\lambda_G(\beta_t)\|^{1/n}.$$

- The convolution operator $\lambda_G(\beta_t)$ in the regular representation obey a remarkable convolution inequality, known as *the Kunze–Stein phenomenon*: for $\beta \in L^r(G)$ with $1 \le r < 2$ and $f \in L^2(G)$,

$$\|\beta * f\|_2 \le c_r \|\beta\|_r \|f\|_2$$

with uniform $c_r > 0$. This convolution inequality for $\mathrm{SL}_2(\mathbb{R})$ is due to Kunze and Stein [KS60] and in general to Cowling [C78]. The Kunze–Stein phenomenon therefore implies that

$$\|\lambda_G(\beta_t)\| \le c_r \left\| \frac{\chi_{B_t}}{|B_t|} \right\|_r = c_r |B_t|^{-1+1/r}.$$

This leads to an estimate for $\left\| \pi_X(\beta_t)|_{L^2_0(X)} \right\|$, and using an interpolation argument, one extends this estimate for the action of G on $L^p(X)$, $1 < p < \infty$. Hence, we deduce the following quantitative mean ergodic theorem, originated in [N98]:

Theorem 4.1. *Given a measure-preserving action with a spectral gap of a connected simple Lie group G on a (standard) probability space (X,μ) , for $1 < p < \infty$ there*

exist $c_p, \theta_p > 0$ such that for the averaging operators $\pi_X(\beta_t)$ as in (4.1),

$$\left\|\pi_X(\beta_t)f - \int_X f d\mu\right\|_p \leq C_p \left|B_t\right|^{-\theta_p} \|f\|_p\,, \quad f \in L^p(X).$$

Moreover, the parameters c_p, θ_p depend only on p and the integrability exponent of the representation $\pi_X|_{L^2_0(X)}$.

It is a remarkable phenomenon that the rate of convergence stated in Theorem 4.1 depends only on the measure of the sets B_t. Thus, L^2-convergence with a rate in a general ergodic action requires no geometric, regularity or stability assumptions of the family of sets in question, so that the quantitative mean ergodic theorem is very robust.

The pointwise ergodic theore, however, is considerably more delicate, and for its validity some stability and regularity assumptions on the family B_t are indispensable. We turn now to discuss these issues in greater detail.

Definition 4.2. An increasing family of bounded measurable subset B_t, $t > 0$, of G is called *admissible* if there exists $c > 0$ such that for all $t \geq t_0$ and $\varepsilon \in (0, \varepsilon_0)$,

$$O_\varepsilon \cdot B_t \cdot O_\varepsilon \subset B_{t+c\varepsilon},$$

where $\mathcal{O}_\epsilon$ denotes the ε-neighbourhood of identity with respect to a right invariant metric on G, and

$$|B_{t+\varepsilon}| \leq (1 + c\,\varepsilon)|B_t|.$$

For instance, if G is embedded as a closed subgroup of $\mathrm{GL}_n(\mathbb{R})$ and $\|\cdot\|$ is norm on $\mathrm{M}_n(\mathbb{R})$, one can show that a family

$$B_t = \{g \in G : \,\log\|g\| < t\}$$

is admissible (see [EMS96, Appendix], [GN10, Ch. 7]).

Under this mild regularity property, we also establish a quantitative pointwise ergodic theorem [MNS00, GN10]:

Theorem 4.3. *Consider a measure-preserving action with a spectral gap of a connected simple Lie group G on a (standard) probability space, and let $\pi_X(\beta_t)$ be the averaging operators with respect to an admissible family of sets. Then there exist $c_p, \theta_p > 0$ such that for every $f \in L^p(X)$, $p > 1$, and μ-almost every $x \in X$,*

$$\left|\pi_X(\beta_t)f(x) - \int_X f\,d\mu\right| \leq C_p(f, x)\,|B_t|^{-\theta'_p},$$

where $\|C_p(f,\cdot)\|_p \leq C'_p\,\|f\|_p$.

One can also establish pointwise ergodic theorem for general actions without the spectral gap assumption, but in this case there is no estimate on the rate of convergence. We refer to [GN10] for a detailed account or ergodic theorem for semisimple groups.

4.2. **On spectral methods in ergodic theory.** Before proceeding with the description of the applications of the mean and pointwise quantitative ergodic theorems stated above, let us make several very brief comments on some of the uses that spectral methods have found in ergodic theory.

1. Euclidean groups : singular averages. The methods of classical harmonic analysis on $\mathbb{R}^d$, including singular integrals, square function estimates, restriction theorems, fractional integrals and analytic interpolation have been utilized by several authors to prove ground-breaking pointwise ergodic theorems. Pioneering contribution here were Stein's ergodic theorem for power of a self adjoint operator [St61], and the celebrated spherical maximal theorem in $\mathbb{R}^d$, proved by Stein-Wainger for $d \geq 3$ [SW78] and by Bourgain for $d = 2$ [Bo83]. The spherical maximal inequalities were used to prove a pointwise ergodic theorem for sphere averages, with the case $d \geq 3$ established by Jones [Jo93], and the case $d = 2$ by Lacey [La95]. Ergodic theorems for sphere averages on the Heisenberg group were developed in [NT97].

For actions of the integers, the methods of discrete harmonic analysis and especially estimates of exponential sums were developed to prove pointwise ergodic theorem for sparse sets of integers, such as square and primes, in Bourgain's pioneering work on the subject [Bo89]. For actions of the integer lattices $\mathbb{Z}^d$, ergodic theorems were developed for discrete spheres and other averages by Magyar [Mag02]. We remark that all the ergodic theorems just mentioned depend on the transference principle for the groups involved, namely they reduce the proof of the requisite maximal inequalities to the case of the action by convolution.

2. Free groups : radial averages. The problem of establishing ergodic theorems for general non-amenable groups was raised already half a century ago by Arnol'd and Krylov [AK63]. An important case that has figured prominently in the theory ever since is that of the free (non-Abelian) group. Arnold and Krylov proved an equidistribution theorem for radial averages on dense free subgroups of isometries of the unit sphere $\mathbb{S}^2$ via a spectral argument similar to Weyl's equidistribution theorem on the circle. Guivarc'h has established a mean ergodic theorem for radial averages on the free group, using von-Neumann's original approach via the spectral theorem [G69]. The pointwise ergodic theorem for general actions of the free group was proved in [N94] for L^2-functions, and extended to function in L^p, $p > 1$, in [NS94]. The distinctly non-amenable phenomenon of an ergodic theorem with a quantitative estimate on the rate of convergence was realized by the celebrated Lubotzky-Phillips-Sarnak construction [LPS86, LPS87] of a dense free group of isometries of $\mathbb{S}^2$ which has an optimal spectral gap.

3. Semisimple groups : ball and sphere averages. An important and influential general spectral result on ergodic actions of semisimple Lie group has been the Howe-Moore mixing theorem [HM79], a consequence of the decay of matrix coefficients in unitary representations of these groups. Tempelman (see [T92]) has used the Howe-Moore theorem to prove mean ergodic theorems for averages on semisimple Lie group.

Some of the methods of classical harmonic analysis alluded to above were combined with unitary representation theory of semisimple groups and used extensively in the development of pointwise ergodic theorems for radial averages on semisimple Lie groups and some of their lattices. Pointwise ergodic theorems for radial averages on the free groups were developed in [N94, NS94]. For sphere and ball averages on simple groups of real rank one pointwise ergodic theorems were developed in [N94b, N97, NS97], and for complex groups in [CN01]. Pointwise ergodic theorems

with a rate of convergence to the ergodic mean were developed for radial averages on general semisimple Lie groups in [MNS00] and for more general averages in [N98].

5. Counting lattice points

Let G be a locally compact second countable group, and Γ is a discrete subgroup with finite covolume. We would be interested in determining the asymptotic behaviour of the number of elements of Γ contained in a family of increasing compact domains B_t in G. The problem of this type arises naturally in arithmetic geometry when $\mathrm{X} \subset \mathbb{C}^n$ is an affine algebraic variety, and one is interested in estimating the cardinality of the set of integral points $x \in \mathrm{X}(\mathbb{Z})$ with bounded norm. The techniques discussed in this section can be applied in particular when $\mathrm{G} \subset \mathrm{GL}_n(\mathbb{C})$ is an algebraic group defined over $\mathbb{Q}$, and $\Gamma = \mathrm{G}(\mathbb{Z})$ is its arithmetic subgroup, and more generally to symmetric varieties of G.

The simplest instance of the lattice counting problem is to estimate the number of integral vectors contained in a compact domain $B \subset \mathbb{R}^d$. In this case, it is easy to see that

$$|\mathbb{Z}^d \cap B| = |B| + O(|\partial B|),$$

and typically $\frac{|\partial B|}{|B|} \to 0$ as the volume $|B| \to \infty$, which leads to an asymptotic formula. The analogous problem for more general groups presents new significant challenges. In particular, for groups with exponential volume growth, $|B|$ is comparable to $|\partial B|$, and the fundamental domain for Γ in G might be unbounded. Thus, in order to establish an asymptotic formula for $|\Gamma \cap B|$, one is required to use more sophisticated analytic techniques. A number of methods have been developed to address this problem, which include, in particular:

- spectral expansion of the automorphic kernel [D42, H56, S56, P76, G83, LP03, MW92, DRS93, BMW99, BGM11],
- decay of matrix coefficients [M02, B82, EM93, Ma07, BO12],
- symbolic coding and transfer operator techniques [L89, P95],
- the theory of unipotent flows [EMS96].

We refer to [B02] for an extensive survey of the lattice counting problem.

Our approach to the lattice point counting problem in the domains B_t is to study the averaging operators defined by B_t in the action of G on the homogeneous space $X = G/\Gamma$ equipped with the invariant probability measure μ. Given a family of domains B_t in G, we consider

$$\pi_X(\beta_t)f(x) = \frac{1}{|B_t|}\int_{B_t} f(g^{-1}x)\, dg, \quad f \in L^2(X).$$

We now turn to demonstrate that a (quantitative) mean ergodic theorem for these operators implies a (quantitative) solution of the lattice point counting problem for the famiy of sets B_t, provided that it satisfies the regularity property of Definition 4.2.

Theorem 5.1. *Let G be a locally compact second countable group, Γ a discrete lattice subgroup and assume that the family of sets B_t is admissible.*

(i) *Assume the averages $\pi_{G/\Gamma}(\beta_t)$ satisfy the mean ergodic theorem:*

$$\left\|\pi_X(\beta_t)f - \int_{G/\Gamma} f\,d\mu\right\|_2 \to 0, \quad f \in L^2(X).$$

Then

$$|\Gamma \cap B_t| \sim \frac{|B_t|}{|G/\Gamma|} \quad as \quad t \to \infty.$$

(ii) *Assume the averages $\pi_X(\beta_t)$ satisfy the quantitative mean ergodic theorem:*

$$\left\|\pi_X(\beta_t)f - \int_X f\,d\mu\right\|_2 \le C_2\,|B_t|^{-\theta}\,\|f\|_2\,, \quad f \in L^2(X,\mu), \tag{5.1}$$

*with $C_2, \theta > 0$. Then for all sufficiently large t,**

$$|\Gamma \cap B_t| = \frac{|B_t|}{|G/\Gamma|} + O\left(|B_t|^{1-\frac{\theta}{\dim(G)+1}+\rho}\right),$$

for every $\rho > 0$, where the implied constants depends only on the parameters in (5.1) and Definition 4.2, and the covolume of Γ.

Let us outline the proof of Theorem 5.1. We first observe that the quantity $|\Gamma\cap B_t|$ can be approximated by suitable averages on space X. Recall that we denote by O_ε the ε-neighbourhood of identity with respect to a right-invariant metric on G, which clearly satisfies $O_\varepsilon^{-1} = O_\varepsilon$. We normalise the invariant measure on G so that Γ has covolume one. Let χ_ε be the characteristic function of this neighbourhood normalised to have integral one and $f_\varepsilon(g\Gamma) = \sum_{\gamma\in\Gamma}\chi_\varepsilon(g\gamma)$. Then $\int_X f\,d\mu = 1$. We claim that for $h \in O_\varepsilon$,

$$\int_{B_{t-c\varepsilon}} f_\varepsilon(g^{-1}h\Gamma)\,dg \le |\Gamma \cap B_t| \le \int_{B_{t+c\varepsilon}} f_\varepsilon(g^{-1}h\Gamma)\,dg. \tag{5.2}$$

Indeed, observe that

$$\int_{B_t} f_\varepsilon(g^{-1}h\Gamma)\,dg = \sum_{\gamma\in\Gamma}\int_{B_t}\chi_\varepsilon(g^{-1}h\gamma)\,dg = \sum_{\gamma\in\Gamma}\frac{|B_t \cap h\gamma O_\varepsilon|}{|O_\varepsilon|}.$$

If $\gamma \in B_{t-c\epsilon}$, then by the admissibility property,

$$h\gamma O_\varepsilon \subset O_\varepsilon B_{t-c\varepsilon} O_\varepsilon \subset B_t,$$

so that

$$|\Gamma \cap B_{t-c\varepsilon}| = \sum_{\gamma\in\Gamma\cap B_{t-c\varepsilon}}\frac{|B_t \cap h\gamma O_\varepsilon|}{|O_\varepsilon|} \le \sum_{\gamma\in\Gamma}\frac{|B_t \cap h\gamma O_\varepsilon|}{|O_\varepsilon|}.$$

On the other hand, if $h\gamma O_\varepsilon \cap B_t \ne \emptyset$, then by the admissibility property,

$$\gamma \in h^{-1}B_t O_\varepsilon \subset O_\varepsilon B_t O_\varepsilon \subset B_{t+c\varepsilon}.$$

Hence,

$$\sum_{\gamma\in\Gamma}\frac{|B_t \cap h\gamma O_\varepsilon|}{|O_\varepsilon|} = \sum_{\gamma\in\Gamma\cap B_{t+c\varepsilon}}\frac{|B_t \cap h\gamma O_\varepsilon|}{|O_\varepsilon|} \le |\Gamma \cap B_{t+c\varepsilon}|.$$

This proves (5.2).

---

* By $\dim(G)$ here we mean the upper local dimension, namely the infimum over $d > 0$ such that $|O_\epsilon| \ge \epsilon^d$ for all sufficiently small $\epsilon > 0$, where O_ε denotes the ε-neighbourhood of identity.

It follows from the mean ergodic theorem that for every $\delta > 0$,

$$\mu(\{h\Gamma \in G/\Gamma : \, |\pi_X(\beta_t) f_\varepsilon(h\Gamma) - 1| > \delta\}) \to 0$$

as $t \to \infty$. In particular, it follows for all sufficiently large t, there exists $h_t \in O_\varepsilon$ such that

$$|\pi_X(\beta_t) f_\varepsilon(h_t\Gamma) - 1| \le \delta. \tag{5.3}$$

Combining this estimate with (5.2), we deduce that

$$|\Gamma \cap B_t| \le (1+\delta)|B_{t+c\varepsilon}| \le (1+\delta)(1+c^2\varepsilon)|B_t| \tag{5.4}$$

and similarly,

$$|\Gamma \cap B_t| \ge (1-\delta)(1-c^2\varepsilon)|B_t| \tag{5.5}$$

for sufficiently large t. Since these estimates hold for arbitrary $\delta, \varepsilon > 0$, this proves the first part of Theorem 5.1.

The proof of the quantitative estimate follows similar ideas, and is equally straightforward. We assume that ε is sufficiently small, and in particular the neighborhoods $\mathcal{O}_\varepsilon\gamma$, $\gamma \in \Gamma$, are disjoint. Then since $|\mathcal{O}_\varepsilon| \ge \varepsilon^n$ with $n > \dim(G)$, we deduce that

$$\|\chi_\varepsilon\|_2^2 = |\mathcal{O}_\varepsilon|^{-1} \le \varepsilon^{-n}.$$

The validity of (5.1) now implies that

$$\mu(\{h\Gamma \in G/\Gamma : \, |\pi_X(\beta_t) f_\varepsilon(h\Gamma) - 1| > \delta\}) \le C_2^2\, \delta^{-2} \varepsilon^{-n} \, |B_t|^{-2\theta} \, .$$

Taking $\delta = 2C_2\varepsilon^{-n}|B_t|^{-\theta}$, we observe that the above measure is strictly less than $\mu(\mathcal{O}_\varepsilon\Gamma) = |\mathcal{O}_\varepsilon|$, so that we can find $h_t \in \mathcal{O}_\varepsilon$ satisfying (5.3) and deduce estimates (5.4)–(5.5) above. Finally, to optimise the estimate, we take $\varepsilon = |B_t|^{-\theta/(n+1)}$ with sufficiently large t. This leads to the second part of the theorem.

Let us note that Theorem 5.1 can be generalized significantly in two important respects. For full details we refer to [GN12a].

(1) The quantitative lattice point count extends to families of sets B_t much more general than admissible ones. It suffices that the families satisfy the following weaker regularity property: there exist $c, a > 0$ such that for all $t \ge t_0$ and $\varepsilon \in (0, \varepsilon_0)$,

$$\left| \bigcup_{u,v \in O_\varepsilon} u B_t v \right| \le (1 + c\,\varepsilon^a) \left| \bigcap_{u,v \in O_\varepsilon} u B_t v \right|,$$

which we call *Hölder well-roundedness*. This extends the solution for example to families of sector averages, families associated naturally with symmetric varieties, and many others.

(2) The quantitative solution to the lattice point counting problem is uniform as the lattice varies in the space of lattice subgroups. All that is needed is that the averaging operators satisfy a uniform norm decay estimate in the space $L^2(G/\Gamma)$, and that G/Γ contain an injective copy of a fixed neighborhood of the identity in G, as Γ varies. In particular, uniformity holds over any set of finite index subgroups which satisfies property (τ). Note that in the ergodic-theoretic approach we have taken, this feature of uniformity is completely obvious from the proof.

In particular, combining the results of this section with Theorem 4.1 and property (τ) for congruence subgroups, we deduce:

Corollary 5.2. *Let* $\mathrm{G} \subset \mathrm{GL}_n(\mathbb{C})$ *be a simply connected simple algebraic group defined over* $\mathbb{Q}$ *and* $\Gamma(\ell) = \{\gamma \in \mathrm{G}(\mathbb{Z}) : \gamma = \mathrm{id} \bmod \ell\}$ *be the family of congruence subgroups. Then for any Hölder well-rounded family of sets* B_t *in* $\mathrm{G}(\mathbb{R})$*, there exists* $\delta > 0$ *such that for sufficiently large* t*,*

$$|\Gamma(\ell) \cap B_t| = \frac{|B_t|}{|\mathrm{G}(\mathbb{R})/\Gamma(\ell)|} + O(|B_t|^{1-\delta}),$$

where the implied constant is independent of ℓ*.*

This estimate will play an indispensable role in the next section.

6. Almost prime points on homogeneous algebraic varieties

In the present section we will describe a solution to the problem of sifting the integral points on certain homogeneous algebraic varieties to obtain almost prime points, following [NS10, GN12c]. As will be explained in due course, the uniformity in the lattice point counting problem established in the previous section will play a crucial role. But let's begin with some remarks on the origins of this problem.

6.1. **Sifting along finite-index subgroups.** The prime number theorem describes the asymptotic distribution of the set $\mathbb{P} = \{2, 3, 5, 7, 11, \ldots\}$ of prime numbers:

$$|\{p \in \mathbb{P} : p \le t\}| \sim \frac{t}{\log t} \quad \text{as } t \to \infty. \tag{6.1}$$

More generally, let us consider an integral irreducible polynomial f. We assume that the leading coefficient of f is positive and $\gcd(f(\mathbb{Z})) = 1$. Is it true that $f(x)$ is prime for infinitely many integers x? The question is answered for linear polynomials by the Dirichlet theorem about primes in arithmetic progressions, but it turns out to be extremely difficult for polynomials of higher degrees. A natural approximation to this problem is the question whether for a fixed $r \ge 1$, the values $f(x)$ belongs to the set $\mathbb{P}_r$ of almost primes infinitely often, where $\mathbb{P}_r$ denotes the set of integers having at most r prime factors. This problem has been extensively studied using sieve methods. For example, it has been shown (see [HR, Sec. 9.5]) that when $r = \deg(f) + 1$, there is a lower estimate of correct order:

$$|\{x \in \mathbb{Z} :\ f(x) \in \mathbb{P}_r, |x| \le t\}| \ge \mathrm{const} \cdot \frac{t}{\log t}. \tag{6.2}$$

Although it is widely believed that under the above necessary conditions on f, the same estimate should holds for the set of primes as well, this is currently out of reach. In particular, it is not known weither there are infinitely many primes of the form $n^2 + 1$ (see [I78] for the best result in this direction).

A far-reaching conjectural generalisation of this problem was proposed by Bourgain, Gamburd and Sarnak (see [S08] and [BGS10]), and termed the saturation problem, as follows. Given

- a Zariski dense subgroup Γ of $\mathrm{GL}_d(\mathbb{Z})$,

- a linear representation $\rho : \mathrm{GL}_d(\mathbb{C}) \to \mathrm{GL}_n(\mathbb{C})$ defined over $\mathbb{Q}$ such that $\rho(\Gamma) \subset \mathrm{GL}_n(\mathbb{Z})$,
- a vector $v \in \mathbb{Z}^n$,
- a polynomial $f \in \mathbb{Z}[x_1, \dots, x_n]$,

is it true that $f(x)$, with $x \in \rho(\Gamma)v$, is r-prime infinitely often, and more generally, is the set of such points x Zariski dense in the Zariski closure of $\rho(\Gamma)v$?

One can also ask for a quantitative version of this question. Setting

$$\mathcal{O}(t) = \{x \in \Gamma v : \|x\| \le t\},$$

is it true that

$$|\{x \in \mathcal{O}(t) : f(x) \in \mathbb{P}_r\}| \ge \text{const} \cdot \frac{|\mathcal{O}(t)|}{\log t} \quad \text{as } t \to \infty, \tag{6.3}$$

in analogy with (6.2)?

It was a profound idea of P. Sarnak that classical sieve methods can be applied in this noncommutative setting provided that one is able to establish an asymptotic estimate for the cardinalities of the sets

$$\mathcal{O}(t, \ell) = \{x \in \Gamma v : \|x\| \le t, \; f(x) = 0 \bmod \ell\} \tag{6.4}$$

as $t \to \infty$. It is crucial for the sieve method to succeed that this estimate is sufficiently uniform over ℓ. Letting

$$\Gamma(\ell) = \{\gamma \in \Gamma : \gamma = \mathrm{id} \bmod \ell\} \tag{6.5}$$

denote the congruence subgroup of level ℓ, it is clear that $f(\gamma\Gamma(\ell)v) = f(\gamma v) \bmod \ell$ for $\gamma \in \Gamma/\Gamma(\ell)$. Thus in order to estimate $|\mathcal{O}(t, \ell)|$, it is sufficient to estimate cardinalities of the sets

$$\{x \in \gamma\Gamma(\ell)v : \|x\| \le t\} \quad \text{for } \gamma \in \Gamma/\Gamma(\ell)\,. \tag{6.6}$$

Let us now consider the important special case of lattice actions on principal homogeneous spaces, where $\mathcal{O} = \Gamma v \subset Gv$, Γ is a lattice in G, and the stability group of v in G is finite. Then estimating (6.6) amounts to nothing other than the lattice point counting problem, specifically in the intersection of norm balls with translates of the orbits of congruence subgroups of Γ. It is crucial for the sieve method to succeed that the solution to the lattice point counting problem in such orbits be uniform over the congruence subgroups. The uniform counting property holds in these sets when Γ is an arithmetic subgroup of a semisimple algebraic group G (for instance, when $\Gamma = \mathrm{SL}_d(\mathbb{Z})$) [NS10]. Thus the required estimates for the sieve method is ultimately derived from the uniform spectral gap property, namely property (τ) discussed in Section 3.1. Indeed, property (τ) gives a uniform spectral estimate for the ball averages acting on the spaces $L^2(\mathrm{G}(\mathbb{R})/\Gamma(\ell))$, which ultimately implies a uniform estimate in the lattice point counting problem.

Establishing a Zariski dense subset of almost prime points on homogeneous varieties requires the more general uniform lattice point counting results of Corollary 5.2. Before stating such a result, let us illustrate how the estimates on $|\mathcal{O}(t, \ell)|$ comes into play on a simple sieve based on the Mobius function. Recall that the

Mobius function $\mu(\ell)$ is equal to $(-1)^r$, if ℓ is a product of r distinct prime factors, and is 0, otherwise, and the Mobius inversion formula

$$\sum_{\ell:\,\ell|n} \mu(\ell) = \begin{cases} 1 \text{ if } n = 1, \\ 0 \text{ if } n > 1. \end{cases}$$

Let P_z be the product of all prime numbers up to z and $S(t,z)$ denotes the cardinality of the set of $x \in \mathcal{O}(t)$ such that $f(x)$ is coprime to P_z. Then

$$\begin{aligned} S(t,z) &= \sum_{k:\gcd(k,P_z)=1} |\mathcal{O}(t) \cap \{f = k\}| \qquad (6.7) \\ &= \sum_{k} \left(\sum_{\ell:\,\ell|\gcd(k,P_z)} \mu(\ell) \right) |\mathcal{O}(t) \cap \{f = k\}| \\ &= \sum_{\ell:\,\ell|P_z} \mu(\ell) \sum_{k:\,\ell|k} |\mathcal{O}(t) \cap \{f = k\}| \\ &= \sum_{\ell:\,\ell|P_z} \mu(\ell) |\mathcal{O}(t,\ell)|. \end{aligned}$$

Now suppose that for some $\kappa, \delta > 0$, we have an estimate

$$|\mathcal{O}(t,\ell)| = \rho(\ell)|\mathcal{O}(t)| + O\left(\ell^\kappa |\mathcal{O}(t)|^{1-\delta}\right) \qquad (6.8)$$

with some uniform $\kappa, \delta > 0$, where

$$\rho(\ell) = \frac{|\{\Gamma v \bmod \ell\} \cap \{f = 0 \bmod \ell\}|}{|\{\Gamma v \bmod \ell\}|}.$$

When z is sufficiently small compared with t, the above computation leads to favourable estimate for $S(t,z)$. However, to derive (6.3), one needs to estimate $S(t,z)$ when $z = t^\alpha$ with some $\alpha > 0$, and the above sieve with the Mobius weights is not sufficiently efficient for our purpose. Fortunately, starting with the works of V. Brunn in 1920's more efficient sieving weights have been developed. We refer to recent surveys [F06, Mo08, Gr10, Mol10] and monographs [FI, G, H, HR, IK] for a comprehensive discussion of sieve methods. These techniques allow to deduce a favourable estimate on $S(t,z)$ with $z = t^\alpha$ with $\alpha > 0$ provided that one knows (6.8) and asymptotic estimates of the coefficients $\rho(\ell)$. The latter estimates can be deduced from the strong approximation property and results on the number of points on varieties over finite fields. Ultimately the main difficulty lies in establishing (6.8), and the crucial initial input is the uniformity in the lattice point counting which follows from property (τ).

More generally, it is possible to apply this method to the case of orbits on symmetric varieties of G. In particular, let us mention the following result [GN12c, Example 1.11], which shows that quadratic surfaces contain many integral points with almost prime coordinates and implies that the set of almost prime points is not contained in any proper algebraic subset. This example is a partial case of Theorem 6.1 below. Let $Q(x) = \sum_{i,j=1}^d a_{ij} x_i x_j$ be a nondegenerate indefinite integral quadratic form in d variables. We assume that $d \geq 4$ (see [LS10] for $d = 3$ case) and denote by Spin_Q the spinor group of Q. We fix $v \in \mathbb{Z}^d$ such that

$n = Q(v) \neq 0$ and consider the orbit $\mathcal{O} = \mathrm{Spin}_Q(\mathbb{Z})v$ which lies in the quadratic surface $\mathrm{X} = \{Q(x) = n\}$. Assuming that the polynomial function f is absolutely irreducible on X and $\gcd(f(x) : x \in \mathcal{O}) = 1$, there exists explicit $r = r(d, \deg(f)) \geq 1$ such that (6.3) holds.

Turning to sifting on general symmetric varieties, recall that Corollary 5.2 establishes uniform asymptotics for the number of elements of $\Gamma(\ell)$ contained in a family of Hölder well-rounded subsets B_t of $\mathrm{G}(\mathbb{R})$. More generally, as noted already in the discussion of almost prime points above, one can also deal of cosets of $\Gamma(\ell)$ and establish the following estimate. For every $\gamma \in \Gamma/\Gamma(\ell)$,

$$|\gamma\Gamma(\ell) \cap B_t| = \frac{|B_t|}{|\mathrm{G}(\mathbb{R})/\Gamma(\ell)|} + O\left(|B_t|^{1-\delta}\right), \tag{6.9}$$

where $\delta > 0$ depends only on the supremum of the integrability exponents in the representations $L_0^2(G/\Gamma(\ell))$, the dimension of G, and the Hölder exponent of B_t's, and the implicit constant is independent of ℓ and $\gamma \in \Gamma/\Gamma(\ell)$. Estimate (6.9) can be used directly for sieving for almost primes on the arithmetic groups Γ itself, or more generally in its orbits in principal homogeneous spaces, and this was carried out in [NS10]. To sieve for almost primes in orbits $\mathcal{O} = \Gamma v$, one may attempt to construct a Hölder well-rounded family $\{B_t\}$ of sets in $\mathrm{G}(\mathbb{R})$ such that the map

$$\Gamma \cap B_t \to \mathcal{O}(t) : \gamma \mapsto \gamma v$$

is onto and has fibers with uniformly bounded cardinalities. Then using (6.9), one would show that there exists $r \geq 1$ such that

$$|\{\gamma \in \Gamma \cap B_t : \ f(\gamma v) \in \mathbb{P}_r\}| \geq \mathrm{const} \cdot \frac{|\Gamma \cap B_t|}{\log t} \quad \text{as } t \to \infty,$$

which implies (6.3). This strategy has been realised in [GN12c] when the stabiliser of v in G is a symmetric subgroup, namely the set of fixed points of an involution. It leads to the following result (see [GN12c, Theorem 1.10]) on the saturation problem for symmetric varieties.

Theorem 6.1. *Let* $\mathrm{G} \subset \mathrm{GL}_n(\mathbb{C})$ *be a connected* $\mathbb{Q}$*-simple simply connected algebraic group,* $\Gamma = \mathrm{G}(\mathbb{Z})$*, and* $v \in \mathbb{Z}^n$ *be such that* $\mathrm{Stab}_{\mathrm{G}}(v)$ *is connected and symmetric. Then given any polynomial* $f : \mathbb{Z}^n \to \mathbb{Z}$ *such the function* $g \mapsto f(gv)$ *is absolutely irreducible on* G *and* $\gcd(f(x) : x \in \mathcal{O}) = 1$*, there exists explicit* $r = r(\mathrm{G}, v, \deg(f)) \geq 1$ *for which the estimate (6.3) holds.*

6.2. **Bounds on integral point satisfying congruence conditions.** Another fundamental question regarding the distribution of prime numbers is the Linnik problem. By Dirichlet's theorem, for every coprime $b, \ell \in \mathbb{N}$ one can find a prime number p satisfying

$$p = b \bmod \ell.$$

Y. Linnik [L44a, L44b] raised the question whether this congruence can be solved effectively, and showed that there exist $\ell_0, \sigma > 0$ such that one can find a prime p satisfying $p \leq \ell^\sigma$ for all $\ell \geq \ell_0$. This problem admits a sweeping generalization to the noncommutative set-up, namely estimating the norm of a minimal prime solution to a system of polynomial equations satisfying a congruence condition. The latter problem is out of reach at this time, but it is possible to give a general

solution to the problem of bounding the norm of an almost prime solution, as follows.

Let $\mathcal{O} \subset \mathbb{Z}^d$ be an orbit of an arithmetic group and $f : \mathcal{O} \to \mathbb{Z}$ a polynomial function. We would like to find a small solution of the congruence

$$f(x) = b \bmod \ell, \quad x \in \mathcal{O}, \tag{6.10}$$

with $f(x)$ being almost prime. Using the standard sieving techniques combined with the estimate (6.9), we prove the following (see [GN12c, Theorem 1.13]):

Theorem 6.2. *Let* $\mathrm{G} \subset \mathrm{GL}_n(\mathbb{C})$ *be a connected* $\mathbb{Q}$*-simple simply connected algebraic group,* $\Gamma = \mathrm{G}(\mathbb{Z})$*, and* $v \in \mathbb{Z}^n$*. Then given a polynomial* $f : \mathbb{Z}^n \to \mathbb{Z}$ *such the function* $g \mapsto f(gv)$ *is absolutely irreducible on* G *and* $\gcd(f(x) : x \in \mathcal{O}) = 1$*, there exist explicit* $r, \ell_0, \sigma > 0$ *such that for every coprime* $b, \ell \in \mathbb{N}$ *satisfying* $\ell \geq \ell_0$ *and* $b \in f(\mathcal{O}) \bmod \ell$*, one can find* x *with* $\|x\| \leq \ell^\sigma$ *satisfying (6.10) such that* $f(x)$ *is* r*-prime.*

The subject of sieving along orbits of noncomutative groups is a rapidly developing area now. In particular, we mentionthe extensive developments regarding Apollonian packings [S07, S09, KO11, BF11] and orbits of "thin" subgroups more generally [BGS06, K09, BGS10, BK10, KO12]. We refer to [K, O] for recent comprehensive surveys.

7. Intrinsic Diophantine approximation on homogeneous varieties

7.1. **Diophantine approximation and ergodic theory.**

In the theory of Diophantine approximation one is interested in finding an efficient rational approximation for vectors $x \in \mathbb{R}^d$. More explicitly, one would like to know for what range of parameters ϵ and R, the system of inequalities

$$\|x - r\| \leq \epsilon \quad \text{and} \quad \mathrm{D}(r) \leq R$$

has a solution $r \in \mathbb{Q}^d$, where $\|\cdot\|$ denotes the maximum norm, and $\mathrm{D}(r)$ denotes the denominator of a rational number r written in reduced form. In this section we consider a more general question regarding intrinsic Diophantine approximation on algebraic sets, namely approximation by rational points on the variety itself. Let

$$\mathrm{X} = \{f_1(x) = \cdots = f_k(x) = 0\} \subset \mathbb{C}^n$$

be an algebraic set defined by a family of polynomial with rational coefficients. For a real points x in $\mathrm{X}(\mathbb{R})$ which is contained in the closure of the set $\mathrm{X}(\mathbb{Q})$ of rational points, we would like to produce an estimate that quantifies density. This amounts to solving the system of inequalities

$$\|x - r\| \leq \epsilon \quad \text{and} \quad \mathrm{D}(r) \leq \epsilon^{-\kappa} \tag{7.1}$$

with $r \in \mathrm{X}(\mathbb{Q})$. The problem of Diophantine approximation on general algebraic varieties by rational points on it was raised in the 1965 survey by Lang [L65]. This question cannot be addressed in full generality because it is not even known how to determine whether $\mathrm{X}(\mathbb{Q})$ is finite or infinite. Here we propose an approach to it when $\mathrm{X}(\mathbb{Q})$ has an additional structure, namely when $\mathrm{X}(\mathbb{Q})$ is equipped with a group action. We note that when X is an elliptic curve (or, more generally, an abelian variety) the problem of Diophantine approximation on X has been studied

by M. Waldschmidt in [W99]. We have considered the problem of Diophantine approximation on an algebraic set X equipped with a transitive action of a semisimple algebraic group G defined over $\mathbb{Q}$. A key new feature of our approach is that it allows approximation by rational points satisfying an arbitrary set of integrality constraints, and not only by the set of all rational points. To simplify our notation for this exposition, we consider the problem of Diophantine approximation by the rational points in $\mathrm{X}(\mathbb{Z}[1/p])$ where p is prime. The problem of Diophantine approximation by points in $\mathrm{X}(\mathbb{Z}[1/p_1,\dots,1/p_s])$ (say), or by $\mathrm{X}(\mathbb{Q})$ can be handled similarly, but uses analysis on adelic spaces which would require much more elaborate notation.

Let us illustrate our general result stated below by the following example, mentioned in the introduction. Let X be a rational two-dimensional ellipsoid such that $\mathrm{X}(\mathbb{Z}[1/p])$ is not discrete in $\mathrm{X}(\mathbb{R})$. Then we prove that

- for almost every $x \in \mathrm{X}(\mathbb{R})$, every $\kappa > 2$, and $\epsilon \in (0, \epsilon_0(x,\kappa))$, the system of inequalities
$$\|x - r\| \le \epsilon \quad \text{and} \quad \mathrm{D}(r) \le \epsilon^{-\kappa} \tag{7.2}$$
has a solution $r \in \mathrm{X}(\mathbb{Z}[1/p])$.
- for every $x \in \mathrm{X}(\mathbb{R})$, every $\kappa > 4$, and $\epsilon \in (0, \epsilon_0(\kappa))$, the system of inequalities
$$\|x - r\| \le \epsilon \quad \text{and} \quad \mathrm{D}(r) \le \epsilon^{-\kappa} \tag{7.3}$$
has a solution $r \in \mathrm{X}(\mathbb{Z}[1/p])$.

We note that the estimate (7.2) is essentially the best possible. Indeed, the number of $r \in \mathrm{X}(\mathbb{Z}[1/p])$ with $\mathrm{D}(r) \le R$ grows as $\mathrm{const} \cdot R$ as $R \to \infty$ and $\dim(\mathrm{X}) = 2$, so that one can deduce from the pigeon-hole principle that $\kappa \ge 2$ in (7.2).

For a general algebraic set X, we set

$$a_p(\mathrm{X}) = \sup_{\text{compact } \Omega \subset \mathrm{X}(\mathbb{R})} \limsup_{R\to\infty} \frac{\log |\{r \in \Omega \cap \mathrm{X}(\mathbb{Z}[1/p]) :\, \mathrm{D}(r) \le R\}|}{\log(R)}.$$

Then again it follows from the pigeon-hole that the exponent $\kappa = \frac{\dim(\mathrm{X})}{a_p(\mathrm{X})}$ in (7.1) is the best possible.

We note that rational ellipsoids constitute the only example for which we are aware of previous results in the literature, as follows. Diophantine approximation by points in $\mathrm{X}(\mathbb{Z}[\frac{1}{p}])$ for $\mathrm{X} = S^2$ and $\mathrm{X} = S^3$ can be deduced from the celebrated construction by Lubotzky-Phillips-Sarnak of a dense subgroup of $\mathrm{SO}(3,\mathbb{R})$ with entries in $\mathbb{Z}[\frac{1}{p}]$ possessing the optimal spectral gap [LPS86, LPS87]. A rate of uniform approximation by all rational points in the spheres S^d, $d \ge 2$ was established by Schmutz [Sch08], using elementary methods based on rational parametrizations of the spheres, the rate being given by $\kappa = 2\log_2(d+1)$. Duke [Du03] has established equidistribution of the set of all rational points on S^2, from which a rate of uniform approximation can be derived, and the same method applies to spheres of any dimension. Finally, we mention the remarkable recent results of Kleinbock and Merrill [KM], which give the best possible results for approximation by all rational points on spheres of any dimension, establishing complete analogs of Dirichlet's and Khinchin's theorems in this case.

Turning now to a general variety, in order to study the density of the set $\mathrm{X}(\mathbb{Z}[1/p])$ in $\mathrm{X}(\mathbb{R})$, we construct a dynamical system which encodes it in some sense. This turns out possible because the set $\mathrm{X}(\mathbb{Z}[1/p])$ is parametrised by the orbits of the group $\mathrm{G}(\mathbb{Z}[1/p])$. It is crucial for the classical theory of Diophantine approximation that $\mathbb{Z}^d$ forms a lattice $\mathbb{R}^d$, and in order to have a similar framework in our setting we consider the topological group $\mathrm{G}(\mathbb{R}) \times \mathrm{G}(\mathbb{Q}_p)$ (where $\mathbb{Q}_p$ denotes the field of p-adic number) equipped with an invariant measure. Then the group $\mathrm{G}(\mathbb{Z}[1/p])$ embeds diagonally in $\mathrm{G}(\mathbb{R}) \times \mathrm{G}(\mathbb{Q}_p)$ as a discrete subgroup with finite covolume. We consider the space

$$Y = (\mathrm{G}(\mathbb{R}) \times \mathrm{G}(\mathbb{Q}_p))/\mathrm{G}(\mathbb{Z}[1/p]).$$

The group $\mathrm{G}(\mathbb{Q}_p)$ acts naturally on this space preserving the finite measure. This is the dynamical system that plays a crucial role in our considerations. It turns out that from the dynamical systems point of view, the original problem of Diophantine approximation corresponds to the *shrinking target property*. Suppose that we have an increasing sequence of compact subsets B_n of $\mathrm{G}(\mathbb{Q}_p)$ and a shrinking sequence of neighbourhoods O_ε in Y. The shrinking target problem asks for what range of parameters n and ϵ the orbits $B_n y$ with $y \in Y$ reach the neighbourhoods O_ϵ. While in the classical theory of dynamical systems, this question is asked for a one-parameter flow, here we are required to answer a similar question for orbits of a large groups $\mathrm{G}(\mathbb{Q}_p)$. The connection between Diophantine approximation in $\mathbb{R}^d$ and the shrinking target property of one-parameter flows have been previously exploited by S. G. Dani, D. Kleinbock, and G. Margulis [D85, D86, D89, K98a, K98b, KM98, K99, KM99]. These developments are surveyed in [K01, M02, K10]. Also in the context of negatively curved manifolds, this connection was explored by S. Hersonsky and F. Paulin [HP01, HP02a, HP02b].

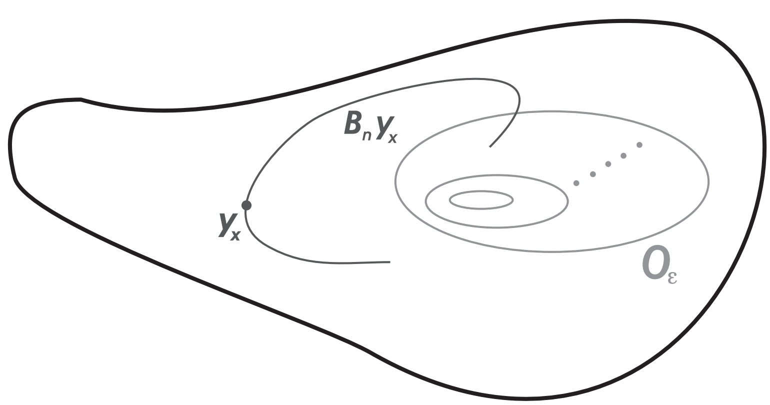

Figure 1. Shrinking target property

7.2. **Diophantine approximation and the duality principle.** To address the problem of Diophantine approximation in our setting we take

$$B_n = \{g \in G(\mathbb{Q}_p) :\ \|g\|_p \le p^n\} \tag{7.4}$$

and construct a shrinking family of neighbourhoods O_ε in Y. An important property of O_ε is that it models the family of neighbourhoods $\|x - r\| \le \epsilon$ in $\mathrm{X}(\mathbb{R})$ and

in particular $|O_\epsilon|$ decays as const $\cdot\, \epsilon^{\dim(X)}$ when $\varepsilon \to 0^+$. We show that every point $x \in X(\mathbb{R})$ is naturally associated a point $y_x \in Y$, and the system of inequalities

$$\|x - r\| \le \epsilon, \quad \mathrm{D}(r) \le p^n$$

has a solution $r \in \mathrm{X}(\mathbb{Z}[1/p])$ provided that

$$B_n^{-1}\, y_x \cap O_\epsilon \ne \emptyset. \tag{7.5}$$

This leads to a useful connection (summarised in Table 1) between the problem of Diophantine approximation on X and the distribution of orbits of Y. In the classical setting an analogous correspondence was observed by S. G. Dani in [D85] and called the Dani correspondence in [KM98]. Since our construction is somewhat similar, we call it the *generalised Dani correspondence*, see Table 1 below.

Diophantine approximation	Dynamics
$\mathrm{X}(\mathbb{Z}[1/p]) \subset \mathrm{X}(\mathbb{R})$	$\mathrm{G}(\mathbb{Q}_p) \curvearrowright Y$
$x \in \mathrm{X}(\mathbb{R})$	$y_x \in Y$
$\|x - r\| \le \epsilon$	$O_\epsilon \subset Y$
$\begin{cases} \|x-r\| \le \epsilon, \\ \mathrm{den}(r) \le p^n \end{cases}$ has a $\mathbb{Z}[1/p]$-solution	$B_n^{-1}\, y_x \cap O_\epsilon \ne \emptyset$

TABLE 1. Generalised Dani correspondence

7.3. **Spectral estimates.** We prove the shrinking target property by analyzing the behavior of suitable averaging operators on the space Y. Let β_n denote the uniform probability measure supported on the set $B_n \subset \mathrm{G}(\mathbb{Q}_p)$. Then we have the sequence of averaging operators

$$\pi_Y(\beta_n) : L^2(Y) \to L^2(Y)$$

defined by

$$\pi_Y(\beta_n)\phi(y) = \int_{\mathrm{G}(\mathbb{Q}_p)} \phi(g^{-1}y)\, d\beta_n(g), \quad \phi \in L^2(Y).$$

Ideally, one hopes that as in the classical von Neumann ergodic theorem, the averages $\pi_Y(\beta_n)\phi$ converge to the space average $\int_Y \phi$ as $n \to \infty$, but this is not always true due to presence of nontrivial unitary characters of $\mathrm{G}(\mathbb{Q}_p)$ in the space $L^2(Y)$. We denote by $L^2_{00}(Y)$ the subspace of $L^2(Y)$ orthogonal to all characters of $\mathrm{G}(\mathbb{R}) \times \mathrm{G}(\mathbb{Q}_p)$. Then one can show that $\pi_Y(\beta_n)\phi \to 0$ in L^2-norm for every

$\phi \in L^2_{00}(Y)$. The rate of convergence here will play a crucial role. This rate is deduce from a variant the spectral gap property, which was discussed in Section 4 and Section 6, and is determined by a suitable integrability exponent. We fix a good maximal compact subgroup U_p of $\mathrm{G}(\mathbb{Q}_p)$. For functions $\phi, \psi \in L^2(Y)$, we denote by $c_{\phi,\psi}(g) = \int_Y \phi(gy)\overline{\psi(y)}\,dy$ the corresponding matrix coefficient. Similarly to (3.8) we define the *spherical integrability exponent*

$$q_p(\mathrm{G}) = \inf\{q > 0 :\ c_{\phi,\psi} \in L^q(\mathrm{G}(\mathbb{Q}_p)) \text{ for all } U_p\text{-inv. } \phi, \psi \in L^2_{00}(Y)\}.$$

It follows from the result of L. Clozel [C03] that the integrability exponent $q_p(\mathrm{G})$ is finite, and we establish the following quantitative mean ergodic theorem: for every sequence B_n of U_p-bi-invariant compact subsets of $\mathrm{G}(\mathbb{Q}_p)$, every $\phi \in L^2_{00}(Y)$ and $\delta > 0$,

$$\|\pi_Y(\beta_n)(\phi)\|_2 \le \mathrm{const}(\delta) \cdot |B_n|^{-\frac{1}{q_p(\mathrm{G})}+\delta}\|\phi\|_2. \tag{7.6}$$

This mean ergodic theorem is vital for establishing the required shrinking target property. We note that in order to derive optimal (or close to optimal) exponents for Diophantine approximation in (7.1) it is crucial that the error term is controlled by the L^2-norm rather than a Sobolev norm, which compromises the rate. The actual value of the exponent of $|B_n|$ in (7.6) is the crucial parameter which ultimately controls the quality of Diophantine approximation, as elaborated further below. Thus this approach provides a connection between the generalised Ramanujan conjectures and the problem of Diophantine approximation on homogeneous varieties.

Now coming back to the original Diophantine approximation problem, we apply estimate (7.6) to the averages supported on the sets B_n as in (7.4) (or, more precisely, on the U_p-bi-invariant sets $U_pB_nU_p$). We also estimate the behaviour of these averages on the orthogonal complement of $L^2_{00}(Y)$ in $L^2(Y)$. Since the orthogonal complement is finite-dimensional, this does not present a particular challenge. Assuming that

$$|O_\epsilon| \ge \mathrm{const} \cdot |B_n|^{-\frac{2}{q_p(G)}+\delta}$$

with fixed $\delta > 0$, we use (7.6) to produce an upper estimate on the measure of the set $y \in Y$ for which $B_n^{-1}y \cap O_\varepsilon = \emptyset$. Then using a Borel–Cantelli-type argument, we deduce a shrinking target property that holds for almost every $y \in Y$. This implies a Diophantine approximation result via the generalised Dani correspondence (Table 1). If we impose a stronger assumption that

$$|O_\epsilon| \ge \mathrm{const} \cdot |B_n|^{-\frac{1}{q_p(G)}+\delta}$$

with some fixed $\delta > 0$, then the above argument can be modified to show that $B_n^{-1}y \cap O_\varepsilon = \emptyset$ for all $y \in Y$ provided that ε is sufficiently small. This produces a Diophantine approximation result which is valid for all points in $\overline{\mathrm{X}(\mathbb{Z}[1/p])}$.

This completes the outline of the following result, which is part of a more general theorem stated in [GGN12, Theorem 1.6]:

Theorem 7.1. *Let* X *be an algebraic set defined over* $\mathbb{Q}$ *equipped with a transitive action of connected almost simple algebraic group* G *defined over* $\mathbb{Q}$*. We assume that* $\mathrm{X}(\mathbb{Z}[1/p])$ *is not discrete in* $\mathrm{X}(\mathbb{R})$†*. Then*

†One can show that under these assumtions $\overline{\mathrm{X}(\mathbb{Z}[1/p])}$ is open in $\mathrm{X}(\mathbb{R})$.

- *for almost every* $x \in \overline{\mathrm{X}(\mathbb{Z}[1/p])} \subset \mathrm{X}(\mathbb{R})$*, every* $\delta > 0$*, and* $\epsilon \in (0, \epsilon_0(x,\delta))$*, the system of inequalities*
$$\|x - r\| \le \epsilon \quad \textit{and} \quad \mathrm{D}(r) \le \left(\epsilon^{-\frac{\dim(\mathrm{X})}{a_p(\mathrm{G})} - \delta}\right)^{q_p(\mathrm{G})/2} \tag{7.7}$$
has a solution $r \in \mathrm{X}(\mathbb{Z}[1/p])$*.*
- *for every* $x \in \mathrm{X}(\mathbb{R})$*, every* $\delta > 0$*, and* $\epsilon \in (0, \epsilon_0(\delta))$*, the system of inequalities*
$$\|x - r\| \le \epsilon \quad \textit{and} \quad \mathrm{D}(r) \le \left(\epsilon^{-\frac{\dim(\mathrm{X})}{a_p(\mathrm{G})} - \delta}\right)^{q_p(\mathrm{G})}$$
has a solution $r \in \mathrm{X}(\mathbb{Z}[1/p])$*.*

In the paper [GGN] (joint with A. Ghosh), we also prove analogues of Khnichin's and Jarnik's theorems for Diophantine approximation on homogeneous varieties.

The quality of Diophantine approximation in Theorem 7.1 is controlled by the integrability exponent $q_p(\mathrm{G})$. When $q_p(\mathrm{G}) = 2$, then estimate (7.7) gives the best possible result up to arbitrary small $\delta > 0$. Coming back to the example (7.2), we note that in this case G is the orthogonal group and $q_p(\mathrm{G}) = 2$, which follows from the work of P. Deligne [D71] on the classical Ramanujan conjecture (see [C] for a comprehensive treatment) combined with the Jacquet–Langlands correspondence [JL70]. This topic is surveyed in [R].

In conclusion we note the intriguing fact that conversely, Diophantine approximation considerations can provide insights in the direction of the generalised Ramanujan conjectures. In particular, one can deduce estimates on the integrability exponents. For instance, analysing the quality of Diophantine approximation in $\mathbb{R}^d$ by the rational points in $\mathbb{Z}[1/p]^d$, one can deduce that
$$q_p(\mathrm{SL}_d) \ge 2(d-1).$$
We refer to [GGN12, Corollary 1.8] for a more general result in this direction.

8. Ergodic theorems for lattice subgroups

8.1. **Ergodic theorems and equidistribution.** We turn to discuss ergodic theorems for countable groups such as $\Gamma = \mathrm{SL}_n(\mathbb{Z})$ and more generally for arithmetic subgroups in semisimple algebraic groups. When G is a connected semisimple group, we have approached this problem via spectral methods, and detailed information about the unitary representation theory of G was essential. However, for discrete lattice subgroups this approach is not feasible, since the irreducible unitary representations of a lattice cannot be classified (in a Borel manner), and a general unitary representation of a lattice cannot be represented uniquely as a direct integral of irreducible ones. Thus non-amenable discrete groups appear as the hardest case in ergodic theory because they possess neither an asymptotically invariant family of sets, nor a usable spectral theory, and so other approaches must be developed.

Although our approach can be applied to lattices in general groups, for the sake of keeping our exposition simple we will present it for lattices Γ in connected simple Lie groups G. For more general formulations, we refer to [GN10]. We take a family of increasing compact domains B_t in G, set $\Gamma_t = \Gamma \cap B_t$, and denote by

λ_t the uniform probability measures supported on the sets Γ_t. Given an arbitrary measure-preserving ergodic action of Γ on a (standard) probability space (X,μ), we consider the corresponding averaging operators

$$\pi_X(\lambda_t)f(x)=\frac{1}{|\Gamma_t|}\sum_{\gamma\in\Gamma_t} f(\gamma^{-1}x),\quad f\in L^p(X).$$

In order for the sets Γ_t to behave in a regular manner, we assume that the domains B_t are admissible in the sense of Definition 4.2. This mild regularity assumption turns out to be sufficient to establish general quantitative ergodic theorems:

Theorem 8.1. *Assume that the action of* Γ *on* (X,μ) *has a spectral gap. Then the following holds.*

- Mean ergodic theorem : *For every* $f\in L^p(X)$ *with* $1<p<\infty$,
$$\left\|\pi_X(\lambda_t)f-\int_X f\,d\mu\right\|_p\le C_p|\Gamma_t|^{-\theta_p}\,\|f\|_p$$
with uniform $C_p,\theta_p>0$, *independent of* f.
- Pointwise ergodic theorem: *For every* $f\in L^p(X)$ *with* $p>1$ *and for* μ-*almost every* x,
$$\left|\pi_X(\lambda_t)f-\int_X f\,d\mu\right|\le C_p(f,x)|\Gamma_t|^{-\theta'_p}$$
with uniform $\theta'_p>0$, *independent of* f *and* x.

We emphasize that this result holds for all Γ-actions which have a spectral gap, and furthermore, if Γ has property T of Kazhdan, the parameters C_p, θ_p and θ'_p are independent of the action. The only connection to the original embedding of Γ in the group G is in the definition of the sets Γ_t. We remark that an L^s-norm bound for $C_p(f,\cdot)$ (for a suitable s depending on p) holds. Finally, when the action of Γ does not have a spectral gap, we also establish mean and pointwise ergodic theorem (but without a rate of convergence, of course). We refer to [GN10] for details.

Theorem 8.1 demonstrates a remarkable contrast with the ergodic theory of amenable groups where no rate of convergence can be established, as discussed in Section 3. Let us give an example to illustrate Theorem 8.1:

- Consider the action $\Gamma=\mathrm{SL}_n(\mathbb{Z})$ on the torus $\mathbb{T}^n=\mathbb{R}^n/\mathbb{Z}^n$ and set
$$\Gamma_t=\{\gamma\in\Gamma:\,\log\|\gamma\|\le t\}.$$
Then for every $f\in L^p(\mathbb{T}^n)$ with $p>1$ and almost every $x\in\mathbb{T}^n$,
$$\left|\frac{1}{|\Gamma_t|}\sum_{\gamma\in\Gamma_t} f(\gamma^{-1}x)-\int_{\mathbb{T}^n} f\,d\mu\right|\le C_p(f,x)|\Gamma_t|^{-\theta'_p}$$
with uniform $\theta'_p>0$.

The orbit structure in the above example is quite complicated. The torus contains the dense set of rational points which consists of finite Γ-orbits, but the orbit of every irrational point is dense. In particular, the size of the estimator $C_p(f,x)$ depends very sensitively on Diophantine properties of x, and it would be interesting

to make this dependence more explicit. In the setting of random walks on the torus, a similar problem was recently investigated in [BFLM11].

In the case of isometric actions of lattices, we derive the following quantitative equidistribution result, which again has no classical amenable analogue:

Corollary 8.2. *Let Γ act on isometrically on a compact Riemannian manifold X equipped with a smooth measure μ of full support. If the action has spectral gap, than for every Hölder function $f \in C^a(X)$, and for every point x in the manifold*

$$\frac{1}{|\Gamma_t|}\sum_{\gamma\in\Gamma_t} f(\gamma^{-1}x) = \int_X f\,d\mu + O\left(|\Gamma_t|^{-\theta_a}\|f\|_{C^a}\right)$$

with uniform $\theta_a > 0$.

8.2. **Method of the proof of the ergodic theorems.** Our approach to proving the ergodic theorems for the averages $\pi_X(\lambda_t)$ is based on the idea of inducing the action of Γ to obtain an action to G on a larger space $Y = (G\times X)/\Gamma$. The latter space has the structure of a fiber bundle over the space G/Γ with fibers isomorphic to X. We then reduce the ergodic theorems for the averaging operators $\pi_X(\lambda_t)$, supported on $\Gamma\cap B_t$ and acting on $L^p(X)$, to the ergodic theorems for the averaging operators $\pi_Y(\beta_t)$, supported on the domains B_t and acting on $L^p(Y)$. This involves developing a series of approximation arguments which constitute generalisations of the argument used in the solution of the lattice point counting problem in Section 5. The essence of the matter is to approximate

$$\pi_X(\lambda_t)\phi(x) = \frac{1}{|\Gamma\cap B_t|}\sum_{\gamma\in\Gamma\cap B_t}\phi(\gamma^{-1}x), \quad \phi\in L^p(X),$$

by

$$\pi_Y(\beta_{t\pm c\varepsilon})f_\varepsilon(y) = \frac{1}{|B_{t\pm c\varepsilon}|}\int_{B_{t\pm c\varepsilon}} f_\varepsilon(g^{-1}y)\,dg.$$

with suitably chosen $f_\varepsilon\in L^p(Y)$. The link between the two expression is given by setting $y = (h,x)\Gamma$ and

$$f_\varepsilon((h,x)\Gamma) = \sum_{\gamma\in\Gamma}\chi_\varepsilon(h\gamma)\phi(\gamma^{-1}x),$$

where χ_ε is the normalized characteristic function of an identity neighborhood O_ε.

The ergodic averages $\pi_Y(\beta_{t\pm c\varepsilon})f_\varepsilon$ can then be rewritten in full as

$$\sum_{\gamma\in\Gamma}\left(\frac{1}{|B_{t\pm c\varepsilon}|}\int_{B_{t\pm c\varepsilon}}\chi_\varepsilon(g^{-1}h\gamma)\,dg\right)\phi(\gamma^{-1}x),$$

We would like the expression in parentheses to be equal to 1 when $\gamma\in\Gamma\cap B_{t-c\varepsilon}$ and equal to 0 when $\gamma\notin\Gamma\cap B_{t+c\varepsilon}$, in order to be able to compare it to $\pi_X(\lambda_t)\phi$. These favourable lower and upper estimates depend only on the regularity properties of the sets B_t, namely, on the admissibility property. Indeed, the lower bound arises since if $\chi_\varepsilon(g^{-1}h\gamma)\neq 0$ and $g\in B_{t-c\varepsilon}$, then $\gamma\in B_t$. The upper bound holds since for $\gamma\in\Gamma\cap B_t$, the support of function $g\mapsto\chi_\varepsilon(g^{-1}h\gamma)$ is contained in $B_{t+c\varepsilon}$.

A fundamental point in the completion of the proof of the ergodic theorems is played by an *invariance principle*, namely by the fact that for any given $f\in L^p(Y)$,

the pointwise ergodic theorem for the averages $\pi_Y(\beta_t)$ holds for a set of points which contain a G-invariant set of full measure. Since the space $Y = (G \times X)/\Gamma$ is a G-equivariant bundle over the G-transitive space G/Γ, this implies that for *every point* $h\Gamma$, the set of points $x \in X$ in the fiber above it where the pointwise ergodic theorem holds is conull in X. This allows us to deduce that the set of points where $\pi_X(\lambda_t)\phi(x)$ converges is also a set of full measure in X.

9. Distribution of orbits on algebraic varieties

9.1. **Distribution of orbits in the de-Sitter space.** Let us now turn to discuss the distribution of dense orbits for actions of arithmetic groups on algebraic varieties. These action do not preserve a finite measure (besides the measures supported on periodic points), so that this problem falls in the realm of infinite ergodic theory.

Consider an action of a discrete group Γ on a space X. The distribution of orbits of Γ is encoded by the properties of the averaging operators

$$\mathcal{A}_t\phi(x) = \frac{1}{|\Gamma_t|}\sum_{\gamma\in\Gamma_t}\phi(\gamma^{-1}x) \tag{9.1}$$

defined for an increasing sequence of finite subsets of Γ that exhaust Γ and for a family of test functions ϕ on X. For example, if $\Gamma = \langle\gamma\rangle$ is a cyclic group, and the action preserves a (possibly infinite) measure μ, then as in Section 1 we obtain that

$$\lim_{N\to\infty}\frac{1}{2N+1}\sum_{n=-N}^{N}\phi(\gamma^n x) \to \mathcal{P}\phi(x) \quad \text{in } L^2(X), \tag{9.2}$$

where $\mathcal{P}$ denotes the orthogonal projection from $L^2(X)$ to the subspace of Γ-invariant functions in $L^2(X)$. However, if the measure is infinite and the action is ergodic, then $\mathcal{P} = 0$, and (9.2) contains no information about the distribution of orbits. Initially, one might expect that if we pick a correct normalisation factor such that $\frac{a(N)}{N} \to 0$, then the averages $\frac{1}{a(N)}\sum_{n=-N}^{N}\phi(\gamma^n x)$ exhibit a nontrivial limit. But in fact, J. Aaranson [Aa, §2.4] showed that if the action of the cyclic group $\Gamma = \langle\gamma\rangle$ is ergodic and conservative, then for any normalisation $a(N)$, either for every nonnegative $\phi \in L^1(X)$,

$$\liminf_{N\to\infty}\frac{1}{a(N)}\sum_{n=-N}^{N}\phi(\gamma^n x) = 0 \quad \text{almost everywhere,}$$

or there exists a subsequence N_k such that such that for every nonnegative $\phi \in L^1(X)$, $\phi \neq 0$,

$$\lim_{k\to\infty}\frac{1}{a(N_k)}\sum_{n=-N_k}^{N_k}\phi(\gamma^n x) = \infty \quad \text{almost everywhere.}$$

But while the averages for orbits of $\mathbb{Z}$-actions on infinite measure spaces tend fluctuate wildly, it turns out that orbits of "large" groups behave in a more regular fashion. We will presently demonstrate this point by computing the distribution

of the orbits of the groups of isometries of the d-dimensional de Sitter space dS_d. This space can be realised as a hypersurface

$$x_1^2 + \cdots + x_d^2 - x_{d+1}^2 = 1$$

in $\mathbb{R}^{d+1}$ equipped with the Minkowski metric $ds^2 = dx_1^2 + \cdots + dx_d^2 - dx_{d+1}^2$. The problem of distribution of orbits on the de Sitter space was raised by V. Arnol'd [Ar, 1996-15, 2002-16]. After stating the result, we will explain a general strategy how to study distribution of orbits on homogeneous spaces. Our exposition is based on [GW07, GN]. We also refer to [L99, N02, Ma02, LP03, G03, G04, GM05, Gu10, N10, MW11] for related results.

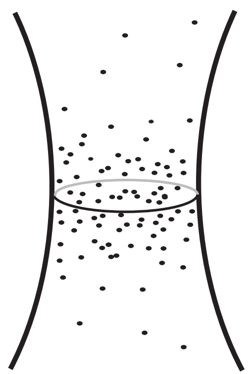

FIGURE 2. The de Sitter space

In the setting of Arnol'd's question regarding the de Sitter space, we prove the following. Let Γ be a discrete subgroup of $\mathrm{Isom}(dS_d)$ with finite covolume. We identify $\mathrm{Isom}(dS_d)$ with the orthogonal group $\mathrm{SO}(d,1)$ and denote by $\|\cdot\|$ the standard Euclidean norm on the space of $(d+1)$-dimensional matrices.

- When $d = 2$, for every $\phi \in L^1(dS^d)$ with compact support
$$\lim_{t\to\infty} \frac{1}{t} \sum_{\gamma\in\Gamma:\, \log\|\gamma\|\le t} \phi(\gamma^{-1}x) = \int_{dS^d} \phi\, d\nu \quad \text{almost everywhere,} \tag{9.3}$$
where ν denotes a (nonzero) invariant measure on dS_d (whose normalization depends on Γ).
- When $d \ge 3$, for every continuous function ϕ with compact support and every $x \in dS^d$ with dense Γ-orbit,
$$\lim_{t\to\infty} \frac{1}{e^{(d-2)t}} \sum_{\gamma\in\Gamma:\, \log\|\gamma\|\le t} \phi(\gamma^{-1}x) = \int_{dS^d} \phi\, d\nu_x, \tag{9.4}$$
where the limit measure is $d\nu_x(y) = \frac{d\nu(y)}{(1+\|x\|^2)^{(d-2)/2}(1+\|y\|^2)^{(d-2)/2}}$.

  Moreover, if in addition ϕ is subanalytic and Sobolev, then for almost every $x \in dS^d$,
$$\frac{1}{e^{(d-2)t}} \sum_{\gamma\in\Gamma:\, \log\|\gamma\|\le t} \phi(\gamma^{-1}x) = \int_{dS^d} \phi(y)\, d\nu_x(y) + O_{\phi,x}(e^{-\delta t}) \tag{9.5}$$
with $\delta = \delta(\phi, x) > 0$.

The above equidistribution results have several remarkable features. In the case of dimension 2, while the cardinality of the set $\{\gamma\in\Gamma:\ \log\|\gamma\|\le t\}$ grows exponentially as $\mathrm{const}\cdot e^t$, it turns out that only for a polynomial number of points γ does the orbit points γx in (9.3) come back to a compact set. Nonetheless, the small fraction of points returning by time t becomes equidistributed in the compact set, with respect to the invariant measure as $t\to\infty$.

When $d\ge 3$, while the cardinality of the set $\{\gamma\in\Gamma:\ \log\|\gamma\|\le t\}$ grows as $\mathrm{const}\cdot e^{(d-1)t}$, only an exponentially small fraction of them satisfy that γx returns to a compact set. The set of returning points does become equidistributied in the compact set, but this time the limiting measures ν_x are not invariant under Γ, and furthermore depend nontrivially on x. The measures ν_x, $x\in X$, on the pseudo-Riemannian space X should be considered as analogues of the Patterson–Sullivan measures in this context.

9.2. **The duality principle on homogeneous spaces.** The techniques developed in [GN] apply in the following general setting. Let $G\subset \mathrm{SL}_d(\mathbb{R})$ be connected and of finite index in an algebraic group, and Γ a discrete subgroup of G with finite covolume. Let X be an algebraic homogeneous space of G equipped with a smooth measure on which Γ acts ergodically. We fix a proper homogeneous polynomial $p:\mathrm{M}_d(\mathbb{R})\to[0,\infty)$ and consider the sets

$$\Gamma_t=\{\gamma\in\Gamma:\ \log p(\gamma)\le t\}.$$

Our goal is to describe the asymptotic distribution of orbits of Γ, or in other words the asymptotic behaviour of the sums $\sum_{\gamma\in\Gamma_t}\phi(\gamma^{-1}x)$ as $t\to\infty$ for a sufficiently rich collection of functions ϕ on X with compact support. In order to find the right normalisation for the sum $\sum_{\gamma\in\Gamma_t}\phi(\gamma^{-1}x)$, one needs to compute what proportion of points in the orbit returns to compact subsets $\Omega\subset X$. Given a compact subset Ω of X with non-empty interior and $x\in X$, we set

$$a=\limsup_{t\to\infty}\frac{\log|\Gamma_t^{-1}x\cap\Omega|}{t}.$$

One can check that this quantity is independent of the choices of Ω and x. We will distinguish two cases according to whether $a=0$, i.e., the return rates are at most subexponentially, and $a>0$, i.e., the return rates are exponential.

Theorem 9.1 (polynomial return rates)**.** *Assume that $a=0$. Then there exists $b\in\mathbb{Z}_{>0}$ such that the averaging operator*

$$\mathcal{A}_t\phi(x)=\frac{1}{t^b}\sum_{\gamma\in\Gamma_t}\phi(\gamma^{-1}x)$$

satisfy

- (strong maximal inequality) *For every $p>1$, a compact $D\subset X$, and $\phi\in L^p(D)$,*

$$\left\|\sup_{t\ge t_0}|\mathcal{A}_t\phi|\right\|_{L^p(D)}\le \mathrm{const}(p,D)\cdot\|\phi\|_{L^p(D)}$$

- (pointwise ergodic theorem) *For every* $\phi \in L^1(X)$ *with compact support,*

$$\lim_{t\to\infty} \mathcal{A}_t\phi(x) = \int_X \phi\, d\nu \quad \textit{almost everywhere,}$$

where ν *is a (nonzero)* G*-invariant measure on* X.

When $a > 0$, we prove analogous results for functions lying in a suitable Sobolev space $L^p_l(X)$. In this case it is also possible to establish rates of convergences.

Theorem 9.2 (exponential return rates)**.** *Assume that* $a > 0$. *Then there exist* $b \in \mathbb{Z}_{\geq 0}$ *and* $l \in \mathbb{Z}_{\geq 0}$ *such that the averaging operator*

$$\mathcal{A}_t\phi(x) = \frac{1}{e^{at}t^b} \sum_{\gamma\in\Gamma_t} \phi(\gamma^{-1}\cdot x)$$

satisfies

- (strong maximal inequality) *For every* $p > 1$, *a compact* $D \subset X$, *and a nonnegative* $\phi \in L^p_l(D)$,

$$\left\| \sum_{t\geq t_0} |\mathcal{A}_t\phi| \right\|_{L^p(D)} \leq \mathrm{const}(p, D)\cdot \|\phi\|_{L^p_l(D)}$$

- (pointwise ergodic theorem) *For every* $p > 1$ *and a nonnegative bounded* $\phi \in L^p_l(X)$ *with compact support,*

$$\lim_{t\to\infty} \mathcal{A}_t\phi(x) = \int_X \phi\, d\nu_x \quad \textit{almost everywhere,} \tag{9.6}$$

where ν_x, $x \in X$, *is a family of absolutely continuous measures on* X *with positive densities.*

- (quantitative pointwise ergodic theorem) *For every* $p > 1$ *and a nonnegative continuous subanalytic* $\phi \in L^p_l(X)$ *with compact support,*

$$\mathcal{A}_t\phi(x) = \int_X \phi\, d\nu_x + \sum_{i=1}^{b} c_i(x,\phi)t^{-i} + O_{\phi,x}\left(e^{-\delta t}\right) \quad \textit{almost everywhere} \tag{9.7}$$

with $\delta = \delta(x, \phi) > 0$.

If in Theorem 9.2 we additionally assume that the stabiliser of a point x in G is semisimple, then the stated results can be further improved. One can replace the Sobolev norms by L^p norms (see [GN, Theorem 1.3]). Moreover, one can show (see [GW07]) for continuous functions ϕ with compact support, the asymptotic formula (9.6) holds for all $x \in X$ whose Γ-orbit in X is dense.

9.3. **Equidistribution in algebraic number fields.** It seems unlikely that in the generality of Theorem 9.2, the asymptotic expansion (9.7) can be proved for an explicit set of $x \in X$ of positive measure, but there is an important case when this can be achieved. Assuming that the action of Γ on X preserves a Riemannian metric, one can show (see [GN12b]) that all Γ-orbits in X become equidistributed with a rate. The observation that all orbits for isometric actions on compact spaces are equidistributed goes back to the work of Y. Guivarc'h [G69], and in [GN12b] we have developed a quantitative version of this argument which also applies to

infinite-measure setting. For instance, fix a non-square integer $d > 0$ and let us consider the action of $\Gamma = \mathrm{SL}_2(\mathbb{Z}[\sqrt{d}])$ on the upper half plane $\mathbb{H}^2$ by fractional linear transformations. We set

$$\Gamma_t = \{\gamma \in \Gamma : \log(\|\gamma\|^2 + \|\bar{\gamma}\|^2) \le t\}$$

where $\|\cdot\|$ is the standard Euclidean norm, and $\bar{\gamma}$ denotes the Galois involution of the field $\mathbb{Q}(\sqrt{d})$. Since Γ embeds diagonally in $G = \mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R})$ as an irreducible lattice, and $\mathbb{H}^2$ is a homogeneous space of the group G, Theorem 9.2 applies in this case and gives information about the asymptotic distribution of orbits Γx for almost every $x \in X$, and the method of [GN12b, Theorem 1.5] leads to the following equidistribution result: there exists $\delta > 0$ such that given a compact subset $D \subset \mathbb{H}^2$, for every $x \in D$ and every $\phi \in C^1(\mathbb{H}^2)$ with support contained in D,

$$\frac{1}{e^t} \sum_{\gamma \in \Gamma_t} \phi(x\gamma) = \int_{\mathbb{H}^2} \phi(z)\, d\nu(z) + O_D(e^{-\delta t}),$$

where ν is a (nonzero) G-invariant measure on $\mathbb{H}^2$.

9.4. **Ideals of the proof.** In conclusion we indicate some ideas that lie behind the proof of Theorem 9.1 and 9.2. The argument can be divided into two main steps:

(I) compare the asymptotic behaviour of discrete averages

$$\sum_{\gamma \in \Gamma:\, \log p(\gamma) \le t} \phi(\gamma^{-1} x) \tag{9.8}$$

with the asymptotic behavior of continuous averages

$$\int_{g \in G:\, \log p(g) \le t} \phi(g^{-1} x)\, dm_G(g), \tag{9.9}$$

where m_G is an invariant measure on G.

(II) establish the asymptotics for the continuous averages (9.9).

To achieve step (I), we think of G as a fiber bundle over X, and show that the fibers of this bundle become equidistributed on the space $\Gamma\backslash G$. More explicitly, we identify X with the homogeneous space G/H where H is a subgroup of G. For $u, v \in G$, we set

$$H_t[u, v] = \{h \in H : \log p(uhv) \le t\}.$$

Then we show that the sum (9.8) can be approximated by a finite linear combination of integrals

$$\int_{H_{t_i}[u_i, v_i]} f_i(y_i h)\, dm_H(h)$$

with suitably chosen $t_i \approx t$, $u_i, v_i \in G$, $f_i : \Gamma\backslash G \to \mathbb{R}$, and $y_i \in \Gamma\backslash G$, where m_H is a left invariant measure on H. The asymptotic behaviour of these integrals can be analysed using either representation-theoretic techniques as in [GN] or the theory of unipotent flows as in [GW07]. In both cases we conclude that as $t \to \infty$,

$$\frac{1}{|H_t[u, v]|} \int_{H_t[u, v]} f(yh) dm_H(h) \to \int_{\Gamma\backslash G} f\, dm_{\Gamma\backslash G},$$

where $m_{\Gamma\backslash G}$ denotes the normalised invariant measure on $\Gamma\backslash G$. Finally, this allows to conclude that the above linear combination is approximately equal to the integral (9.9), which completes step (I).

In step (II), we establish an asymptotic formula for integral

$$v(t) = \int_{g\in G:\,\log p(g)\le t} \phi(g^{-1}x)\,dm(g)$$

where ϕ is a continuous subanalytic function with compact support. For this, we consider its transform

$$f(s) = \int_0^\infty t^{-s} v(\log t)\,dt$$

which converges when $\mathrm{Re}(s)$ is sufficiently large. Using a suitable version of resolution of singularities as in [P94] we show that $f(s)$ has a meromorphic continuation beyond the first pole. Then the asymptotic formula for $v(t)$ follows via a Tauberian argument.

10. Acknowledgements

Part of this paper was written in Spring 2011 when A.G. was visiting École Polytechnique Fédérale de Lausanne during the programme "Group Actions in Number Theory", and he would like to thank Emmanuel Kowalski, Philippe Michel, and the Centre Interfacultaire Bernoulli for hospitality.

Both authors would like to express their gratitude to the Ergodic Theory Group at the Fédération Denis Poisson for the opportunity to explain the present work in the lecture series "Théorie ergodique des actions de groupes" held at the University of Tours in April, 2011. In particular we would like to thank the organizers, Claire Anantharaman-Delaroche, Jean-Philippe Anker, and Emmanuel Lesigne.

A.G. also would like to thank for hospitality Manfred Einsiedler and the Institute for Mathematical Research at ETH, Zürich during his visit in Autumn 2012 when the work on this survey was completed.

A.G. was supported in part by EPSRC, ERC, and RCUK, and A.N. was supported by an ISF grant.

References

[Aa] J. Aaronson, An introduction to infinite ergodic theory. Mathematical Surveys and Monographs, 50. American Mathematical Society, Providence, RI, 1997.

[AAB] C. Anantharaman, J.-P. Anker, M. Babillot, A. Bonami, B. Demange, S. Grellier, F. Havard, P. Jaming, E. Lesigne, P. Maheux, J.-P. Otal, B. Schapira, J.-P. Schreiber, Théorèmes ergodiques pour les actions de groupes. Monographies de L'Enseignement Mathématique 41. L'Enseignement Mathématique, Geneva, 2010.

[Ar] V. Arnol'd, Arnold's problems. Springer-Verlag, Berlin, 2004.

[AK63] V. Arnol'd and A. Krylov, Uniform distribution of points on a sphere and certain ergodic properties of solutions of linear ordinary differential equations in a complex domain. Dokl. Akad. Nauk SSSR 148 (1963), 9–12.

[B02] M. Babillot, Points entiers et groupes discrets: de l'analyse aux systmes dynamiques. Panor. Synthèses, 13, Rigidité, groupe fondamental et dynamique, 1–119, Soc. Math. France, Paris, 2002.

[B82] H.-J. Bartels, Nichteuklidische Gitterpunktprobleme und Gleichverteilung in linearen algebraischen Gruppen. Comment. Math. Helv. 57 (1982), no. 1, 158–172.

[BHV] B. Bekka, P. de la Harpe and A. Valette, Kazhdan's property (T). New Mathematical Monographs, 11. Cambridge University Press, Cambridge, 2008.

[BO12] Y. Benoist and H. Oh, Effective equidistribution of S-integral points on symmetric varieties, Ann. Inst. Fourier 62 (2012), 1889–1942.

[B32] G. D. Birkhoff, Proof of the ergodic theorem, Proc. Nat. Acad. Sci. USA 17 (1931), 656–660.

[BB13] V. Bloomer and F. Brumley, The role of the Ramanujan conjecture in analytic number theory, Bulletin AMS 50 (2013), 267–320.

[BW80] A. Borel and N. Wallach, Continuous cohomology, discrete groups and representation of reductive groups, Annals of Mathematics Studies 94, Princeton University Press, 1980.

[BF11] J. Bourgain and E. Fuchs, A proof of the positive density conjecture for integer Apollonian circle packings, J. Amer. Math. Soc. 24 (2011), 945–967.

[Bo83] J. Bourgain, Averages in the plane over convex curves and maximal operators, J. Anal. Math. 47 (1983), 69–85.

[Bo89] J. Bourgain, Pointwise ergodic theorems for arithmetic sets. With an appendix by the author, H. Furstenberg, Y. Katznelson and D. S. Ornstein. Publ. Math. Inst. Hautes E?tudes Sci. 69 (1989), 5–45.

[BFLM11] J. Bourgain, A. Furman, E. Lindenstrauss, S. Mozes, Stationary measures and equidistribution for orbits of nonabelian semigroups on the torus. J. Amer. Math. Soc. 24 (2011), no. 1, 231–280.

[BGS06] J. Bourgain, A. Gamburd and P. Sarnak, Sieving and expanders. C. R. Math. Acad. Sci. Paris 343 (2006), no. 3, 155–159.

[BGS10] J. Bourgain, A. Gamburd and P. Sarnak, Affine linear sieve, expanders, and sum-product. Invent. Math. 179 (2010), no. 3, 559–644.

[BK10] J. Bourgain and A. Kontorovich, On Representations of Integers in Thin Subgroups of $SL(2, Z)$, Geom. Funct. Anal. 20 (2010), 1144–1174; erratum: Geom. Funct. Anal. 20 (2010), 1548–1549.

[BMW99] R. Bruggeman, R. Miatello and N. Wallach, Resolvent and lattice points on symmetric spaces of strictly negative curvature. Math. Ann. 315 (1999), 617–639.

[BGM11] R. Bruggeman, F. Grunewald and R. Miatello, New lattice point asymptotics for products of upper half-planes. Int. Math. Res. Not. (2011), no. 7, 1510–1559.

[BS91] M. Burger, and P. Sarnak, Ramanujan duals. II. Invent. Math. 106 (1991), no. 1, 1–11.

[C53] A.-P. Calderon, A general ergodic theorem, Ann. of Math. 58 (1953), 182–191.

[CT10] A. Chambert-Loir and Y. Tschinkel, Igusa integrals and volume asymptotics in analytic and adelic geometry, Confluentes Mathematici 2 (2010), no. 3, 351–429.

[C03] L. Clozel, Démonstration de la conjecture τ. Invent. Math. 151 (2003), no. 2, 297–328.

[CW76] R. Coifman and G. Weiss, Transference methods in analysis. Conference Board of the Mathematical Sciences Regional Conference Series in Mathematics, No. 31. American Mathematical Society, Providence, R.I., 1976.

[C] B. Conrad, Modular forms and the Ramanujan conjecture, Cambridge University Press, 2011.

[C78] M. Cowling, The Kunze-Stein phenomenon, Ann. Math. 107 (1978), no. 2, 209–234.

[C79] M. Cowling, Sur les coefficients des représentations unitaires des groupes de Lie simples. Analyse harmonique sur les groupes de Lie (Sém., Nancy-Strasbourg 1976–1978), II, pp. 132–178, Lecture Notes in Math., 739, Springer, Berlin, 1979.

[CN01] M. Cowling and A. Nevo, Uniform estimates for spherical functions on complex semisimple Lie groups. Geom. Funct. Anal. 11 (2001), 900–932.

[D85] S. G. Dani, Divergent trajectories of flows on homogeneous spaces and Diophantine approximation, J. Reine Angew. Math. 359 (1985), 55–89; erratum: J. Reine Angew. Math. 359 (1985), 214.

[D86] S. G. Dani, Bounded orbits of flows on homogeneous spaces. Comment. Math. Helv. 61 (1986), no. 4., 636–660.

[D89] S. G. Dani, On badly approximable numbers, Schmidt games and bounded orbits of flows. Number theory and dynamical systems (York, 1987), 69–86, London Math. Soc. Lecture Note Ser., 134, Cambridge Univ. Press, Cambridge, 1989.

[D71] P. Deligne, Formes modulaires et représentations l-adiques, Séminaire Bourbaki vol. 1968/69 Exposés 347–363, Lecture Notes in Mathematics, 179, Berlin, New York: Springer-Verlag, 1971.

[D42] J. Delsarte, Sur le gitter fuchsien. C. R. Acad. Sci. Paris 214, (1942). 147–179.

[DH] H. Diamond and H. Halberstam, A higher-dimensional sieve method. Cambridge Tracts in Mathematics, 177. Cambridge University Press, Cambridge, 2008.

[Du03] W. Duke, Rational points on the sphere. Rankin memorial issues. Ramanujan J. 7 (2003), no. 1-3, 235–239.

[DRS93] W. Duke, Z. Rudnick, P. Sarnak, Density of integer points on affine homogeneous varieties. Duke Math. J. 71 (1993), no. 1, 143–179.

[EG74] F. Greenleaf and W. Emerson, Group structure and the pointwise ergodic theorem for connected amenable groups. Advances in Math. 14 (1974), 153–172.

[EM93] A. Eskin and C. McMullen, Mixing, counting, and equidistribution in Lie groups. Duke Math. J. 71 (1993), no. 1, 181–209.

[EMS96] A. Eskin, S. Mozes and N. Shah, Unipotent flows and counting lattice points on homogeneous varieties. Ann. of Math. 143 (1996), no. 2, 253–299.

[F06] J. Friedlander, Producing prime numbers via sieve methods. Analytic number theory, 1–49, Lecture Notes in Math., 1891, Springer, Berlin, 2006.

[FI] J. Friedlander and H. Iwaniec, Opera de cribro. American Mathematical Society Colloquium Publications, 57. American Mathematical Society, Providence, RI, 2010.

[GJ78] S. Gelbart and H. Jacquet, A realtion between automorphic representations of $GL(2)$ and $GL(3)$. Ann. Sci. Ecole Norm. Sup. 11 (1978), 471–552.

[GGN12] A. Ghosh, A. Gorodnik and A. Nevo, Diophantine approximation and automorphic spectrum, Int. Math. Research Notices 2012, doi:10.1093/imrn/rns198.

[GGN] A. Ghosh, A. Gorodnik and A. Nevo, Metric Diophantine approximation on homogeneous varieties, arXiv:1205.4426.

[G83] A. Good, Local analysis of Selberg's trace formula. Lecture Notes in Mathematics, 1040. Springer-Verlag, Berlin, 1983.

[G03] A. Gorodnik, Lattice action on the boundary of $SL(n, R)$. Ergodic Theory Dynam. Systems 23 (2003), no. 6, 1817–1837.

[G04] A. Gorodnik, Uniform distribution of orbits of lattices on spaces of frames. Duke Math. J. 122 (2004), no. 3, 549–589.

[GM05] A. Gorodnik and F. Maucourant, Proximality and equidistribution on the Furstenberg boundary. Geom. Dedicata 113 (2005), 197–213.

[GN10] A. Gorodnik and A. Nevo, The ergodic theory of lattice subgroups, Annals of Mathematics Studies 172, Princeton University Press, 2010.

[GN12a] A. Gorodnik and A. Nevo, Counting lattice points, J. Reine Angew. Math. 663 (2012), 127–176.

[GN12b] A. Gorodnik and A. Nevo, On Arnold's and Kazhdan's equidistribution problems, Ergodic Theory Dynam. Systems 32 (2012), no. 6, 1972–1990.

[GN12c] A. Gorodnik and A. Nevo, Lifting, restricting and sifting integral points on affine homogeneous varieties, Compositio Math. 2012, doi:10.1112/S0010437X12000516.

[GN] A. Gorodnik and A. Nevo, Ergodic theory and the duality principle on homogeneous spaces, arXiv:1205.4413.

[GW07] A. Gorodnik and B. Weiss, Distribution of lattice orbits on homogeneous varieties. Geom. Funct. Anal. 17 (2007), no. 1, 58–115.

[G] G. Greaves, Sieves in number theory. Ergebnisse der Mathematik und ihrer Grenzgebiete 43. Springer-Verlag, Berlin, 2001.

[Gr10] A. Granville, Different approaches to the distribution of primes. Milan J. Math. 78 (2010), no. 1, 65–84.

[Gu10] A. Guilloux, Polynomial dynamic and lattice orbits in S-arithmetic homogeneous spaces. Confluentes Math. 2 (2010), no. 1, 1–35.

[G69] Y. Guivarc'h, Généralisation d'un théorème de von Neumann, C.R. Acad. Sci. Paris 268 (1969), 1020–1023.

[HR] H. Halberstam and H.-E. Richert, Sieve methods. London Mathematical Society Monographs, No. 4. Academic Press, London-New York, 1974.

[H] G. Harman, Prime-detecting sieves. London Mathematical Society Monographs Series, 33. Princeton University Press, Princeton, NJ, 2007.

[HP01] S. Hersonsky and F. Paulin, Hausdorff dimension of Diophantine geodesics in negatively curved manifolds. J. Reine Angew. Math. 539 (2001), 29–43.

[HP02a] S. Hersonsky and F. Paulin, Diophantine approximation in negatively curved manifolds and in the Heisenberg group. Rigidity in dynamics and geometry (Cambridge, 2000), 203–226, Springer, Berlin, 2002.

[HP02b] S. Hersonsky and F. Paulin, Diophantine approximation for negatively curved manifolds. Math. Z. 241 (2002), no. 1, 181–226.

[HM79] R. Howe and C. Moore, Asymptotic properties of unitary representations. J. Funct. Anal. 32 (1979), no. 1, 72–96.

[HT] R. Howe and E.-C. Tan, Nonabelian harmonic analysis. Applications of SL(2,R). Universitext. Springer-Verlag, New York, 1992.

[H56] H. Huber, Über eine neue Klasse automorpher Funktionen und ein Gitterpunktproblem in der hyperbolischen Ebene. I. Comment. Math. Helv. 30 (1956), 20–62.

[I78] H. Iwaniec, Almost primes represented by quadratic polynomials. Invent. Math. 47 (1978), 171–188.

[I96] H. Iwaniec, The lowest eigenvalue of congruence subgroups. Topics in geometry, 203–212, Progr. Nonlinear Differential Equations Appl., 20, Birkhäuser, Boston, 1996.

[IK] H. Iwaniec and E. Kowalski, Analytic number theory. American Mathematical Society Colloquium Publications, 53. American Mathematical Society, Providence, RI, 2004.

[JL70] H. Jacquet and R. Langlands, Automorphic forms on $GL(2)$. Lecture notes in Mathematics 114, Springer-Verlag, Berlin-New York, 1970.

[Jo93] R. Jones, Ergodic averages on spheres. J. Anal. Math. 61 (1993), 29–45.

[JR79] A. del Junco and J. Rosenblatt, Counterexamples in ergodic theory and number theory. Math. Ann. 245 (1979), no. 3, 185–197.

[KW82] Y. Katznelson and B. Weiss, A simple proof of some ergodic theorems. Israel J. Math. 42 (1982), no. 4, 291–296.

[K65] D. Kazhdan, Uniform distribution on a plane. Trudy Moskov. Mat. Obsh. 14 (1965), 299–305.

[K67] D. Kazhdan, On the connection of the dual space of a group with the structure of its closed subgroups. Funkcional. Anal. i Prilozen. 1 (1967), 71–74.

[KS03] H. Kim and P. Sarnak, Refined estimates towards the Ramanujan and Selberg conjectures, appendix in H. Kim, Functoriality for the exterior square of GL(4) and the symmetric fourth of GL(2). J. Amer. Math. Soc. 16 (2003), no. 1, 139–183.

[K98a] D. Kleinbock, Flows on homogeneous spaces and Diophantine properties of matrices. Duke Math. J. 95 (1998), no. 1, 107–124.

[K98b] D. Kleinbock, Bounded orbits conjecture and Diophantine approximation. Lie groups and ergodic theory (Mumbai, 1996), 119–130, Tata Inst. Fund. Res. Stud. Math., 14, Bombay, 1998.

[K99] D. Kleinbock, Badly approximable systems of affine forms. J. Number Theory 79 (1999), no. 1, 83–102.

[K01] D. Kleinbock, Some applications of homogeneous dynamics to number theory. Smooth ergodic theory and its applications (Seattle, WA, 1999), 639–660, Proc. Sympos. Pure Math., 69, Amer. Math. Soc., Providence, RI, 2001.

[K10] D. Kleinbock, Quantitative nondivergence and its Diophantine applications. Homogeneous flows, moduli spaces and arithemtic, 131–153, Clay Math. Proc., 10, Amer. Math. Soc., Providence, RI, 2010.

[KM98] D. Kleinbock and G. Margulis, Flows on homogeneous spaces and Diophantine approximation on manifolds. Ann. of Math. (2) 148 (1998), no. 1, 339–360.

[KM99] D. Kleinbock and G. Margulis, Logarithm laws for flows on homogeneous spaces. Invent. Math. 138 (1999), no. 3, 451–494.

[KM] D. Kleinbock and K. Merrill, Rational approximation on spheres. arXiv:1301.0989.
[K09] A. Kontorovich, The Hyperbolic Lattice Point Count in Infinite Volume with Applications to Sieves, Duke Math J. (149) 2009, no. 1, 1–36.
[K] A. Kontorovich, From Apollonius To Zaremba: Local-Global Phenomena in Thin Orbits. arXiv:1208.5460.
[KO11] A. Kontorovich and H. Oh, Apollonian Circle Packings and Closed Horospheres on Hyperbolic 3-Manifolds, J. Amer. Math. Soc 24, (2011), no. 3, 603–648.
[KO12] A. Kontorovich and H. Oh, Almost Prime Pythagorean Triples in Thin Orbits, J. Reine Angew. Math., 667 (2012), 89–131.
[KS60] R. A. Kunze and E. M. Stein, Uniformly bounded representations and harmonic analysis of the 2×2 real unimodular group, Amer. J. Math. 82 (1960), 1–62.
[K69] B. Kostant. On the existence and irreducibility of certain series of representations. Bull. Amer. Math. Soc. 75 (1969), 627–642.
[La95] M. Lacey, Ergodic averages on circles. J. Anal. Math. 67 (1995), 199–206.
[L89] S. Lalley, Renewal theorems in symbolic dynamics, with applications to geodesic ows, non-Euclidean tessellations and their fractal limits. Acta Math. 163 (1989), no. 1-2, 1–55.
[L65] S. Lang, Report on Diophantine Approximation, Bull. Soc. Math. France 93 (1965), 177–192.
[LP82] P. Lax and R. Phillips, The asymptotic distribution of lattice points in Euclidean and non-Euclidean spaces. J. Funct. Anal. 46 (1982), no. 3, 280–350.
[L01] E. Lindenstrauss, Pointwise theorems for amenable groups, Invent. Math. 146 (2001), no. 2, 259–295.
[L99] F. Ledrappier, Distribution des orbites des réseaux sur le plan réel. C. R. Acad. Sci. Paris Sér. I Math. 329 (1999), no. 1, 61–64.
[LP03] F. Ledrappier and M. Pollicott, Ergodic properties of linear actions of (2×2)-matrices. Duke Math. J. 116 (2003), no. 2, 353–388.
[L95] J.-S. Li, The minimal decay of matrix coefficients for classical groups. Harmonic analysis in China, 146–169, Math. Appl., 327, Kluwer Acad. Publ., Dordrecht, 1995
[LZ96] J.-S. Li and C.-B. Zhu, On the decay of matrix coefficients for exceptional groups. Math. Ann. 305 (1996), no. 2, 249–270.
[L44a] Y. Linnik, On the least prime in an arithmetic progression. I. The basic theorem. Rec. Math. [Mat. Sbornik] N.S. 15(57) (1944), 139–178.
[L44b] Y. Linnik, On the least prime in an arithmetic progression. II. The Deuring-Heilbronn phenomenon. Rec. Math. [Mat. Sbornik] N.S. 15(57) (1944), 347–368.
[LS10] J. Liu and P. Sarnak, Integral points on quadrics in three variables whose coordinates have few prime factors. Israel J. Math. 178 (2010), 393–426.
[Lu12] A. Lubotzky, Expander graphs in pure and applied mathematics. Bull. Amer. Math. Soc. (N.S.) 49 (2012), no. 1, 113–162.
[LPS86] A. Lubotzky, R. Phillips, P. Sarnak, Hecke operators and distributing points on the sphere. I. Frontiers of the mathematical sciences: 1985 (New York, 1985). Comm. Pure Appl. Math. 39 (1986), no. S, suppl., S149–S186.
[LPS87] A. Lubotzky, R. Phillips, P. Sarnak, Hecke operators and distributing points on S^2. II. Comm. Pure Appl. Math. 40 (1987), no. 4, 401–420.
[LRS95] W. Luo, Z. Rudnick and P. Sarnak, On Selberg's eigenvalue conjecture, Geom. Funct. Anal. 5 (1995), 387–401.
[Mag02] A. Magyar, Diophantine equations and ergodic theorems. Amer. J. Math. 124 (2002), 921–953.
[M02] G. Margulis, Diophantine approximation, lattices and flows on homogeneous spaces. A panorama of number theory or the view from Baker's garden (Zürich, 1999), 280–310, Cambridge Univ. Press, Cambridge, 2002.
[M04] G. Margulis, On some aspects of the theory of Anosov systems. With a survey by Richard Sharp: Periodic orbits of hyperbolic flows. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2004.

[MNS00] G. A. Margulis, A. Nevo, E. M. Stein, Analogs of Wiener's ergodictheorems for semisimple Lie groups. II, Duke Math. J. 103 (2000), no. 2, 233–259.

[Ma02] F. Maucourant, Approximation diophantienne, dynamique des chambres de Weyl et répartitions d'orbites de réseaux, PhD Thesis, Université de Lille, 2002.

[Ma07] F. Maucourant, Homogeneous asymptotic limits of Haar measures of semisimple linear groups and their lattices. Duke Math. J. 136 (2007), no. 2, 357–399.

[MW11] F. Maucourant and B. Weiss, Lattice actions on the plane revisited, Geom. Dedicata 157 (2012), 1–21.

[MW92] R. Miatello and N. Wallach, The resolvent of the Laplacian on locally symmetric spaces. J. Differential Geom. 36 (1992), no. 3, 663–698.

[Mol10] R. Mollin, An overview of sieve methods, Int. J. Contemp. Math. Sci. 5 (2010), no. 1–4, 67–80.

[Mo08] Y. Motohashi, An overview of sieve methods and their history, Sugaku Expositions 21 (2008), 1–32.

[N94] A. Nevo, Harmonic analysis and pointwise ergodic theorems for noncommuting transformations. J. Amer. Math. Soc. 7 (1994), no. 4, 875–902.

[N94b] A. Nevo, Pointwise ergodic theorems for radial averages on simple Lie groups I. Duke Math. J. 76 (1994), 113–140.

[N97] A. Nevo, Pointwise ergodic theorems for radial averages on simple Lie groups II. Duke Math. J. 86 (1997), 239–259.

[N98] A. Nevo, Spectral transfer and pointwise ergodic theorems for semi-simple Kazhdan groups. Math. Res. Lett. 5 (1998), no. 3, 305–325.

[N06] A. Nevo, Pointwise ergodic theorems for actions of groups. Handbook of dynamical systems. Vol. 1B, 871–982, Elsevier B. V., Amsterdam, 2006.

[NS10] A. Nevo and P. Sarnak, Prime and almost prime integral points on principal homogeneous spaces, Acta Math. 205 (2010), no. 2, 361–402.

[NS94] A. Nevo and E. Stein, A generalization of Birkhoff's pointwise ergodic theorem. Acta Math. 173 (1994), no. 1, 135–154.

[NS97] A. Nevo and E. M. Stein, Analogs of Wiener's ergodic theorems for semi-simple Lie groups I. Ann. of Math., **145** (1997), pp. 565–595.

[NT97] A. Nevo and S. Thangavelu, Pointwise ergodic theorems for radial averages on the Heisenberg group. Adv. Math. 127 (1997), 307 334.

[N03] M. Neuhauser, Kazhdan constants and matrix coefficients of $Sp(n, R)$. J. Lie Theory 13 (2003), no. 1, 133–154.

[N02] A. Nogueira, Orbit distribution on R^2 under the natural action of $SL(2, Z)$. Indag. Math. (N.S.) 13 (2002), no. 1, 103–124.

[N10] A. Nogueira, Lattice orbit distribution on R^2. Ergodic Theory Dynam. Systems 30 (2010), no. 4, 1201–1214; erratum: Ergodic Theory Dynam. Systems 30 (2010), no. 4, 1215.

[O02] H. Oh, Uniform pointwise bounds for matrix coefficients of unitary representations and applications to Kazhdan constants. Duke Math. J. 113 (2002), no. 1, 133–192.

[O] H. Oh, Harmonic analysis, Ergodic theory and Counting for thin groups. arXiv:1208.4148.

[P76] S. J. Patterson, A lattice-point problem in hyperbolic space. Mathematika 22 (1975), no. 1, 81–88; corrigendum: Mathematika 23 (1976), no. 2, 227.

[P94] A. Parusinski, Subanalytic functions. Trans. Amer. Math. Soc. 344 (1994), no. 2, 583–595.

[P95] M. Pollicott, A symbolic proof of a theorem of Margulis on geodesic arcs on negatively curved manifolds. Amer. J. Math. 117 (1995), no. 2, 289–305.

[R38] F. Riesz, Some mean ergodic theorems, J. Lond. Math. Soc. 1 (1938), 274–278.

[R] J. Rogawski, Modular forms, the Ramanujan conjecture, and the Jacquet-Langlands correspondence; appendix in A. Lubotzky, Discrete groups, expanding graphs and invariant measures. Progress in Mathematics, 125. Birkhaäuser Verlag, Basel, 1994.

[S95] P. Sarnak, Selberg's eigenvalue conjecture. Notices Amer. Math. Soc. 42 (1995), no. 11, 1272–1277.

[S03] P. Sarnak, Spectra of hyperbolic surfaces, Bull. Amer. Math. Soc. 40 (2003), no. 4, 441–478.

[S05] P. Sarnak, Notes on the generalized Ramanujan conjectures. Harmonic analysis, the trace formula, and Shimura varieties, 659–685, Clay Math. Proc., 4, Amer. Math. Soc., Providence, RI, 2005.
[S07] P. Sarnak, Letter to Lagarias on integral Apollonian packings, http://www.math.princeton.edu/sarnak/
[S08] P. Sarnak, Equidistribution and primes, http://www.math.princeton.edu/sarnak/
[S09] P. Sarnak, Integral Apollonian Packings, http://www.math.princeton.edu/sarnak/
[S81] K. Schmidt. Amenability, Kazhdan's property (T), strong ergodicity and invariant means for group actions. Ergod. Theory Dynam. Systemes 1 (1981), 223–236.
[Sch08] E. Schmutz, Rational points on the unit sphere, Cent. Eur. J. Math. 6 (2008), no. 3, 482487.
[S56] A. Selberg, Harmonic analysis and discontinuous groups in weakly symmetric Riemannian spaces with applications to Dirichlet series. J. Indian Math. Soc. (N.S.) 20 (1956), 47–87.
[S65] A. Selberg, On estimation of Fourier coefficients of modular forms, Proc. Sympos. Pure Math., Amer. Math. Soc., vol. VIII, 1965, 1–15.
[St61] E. M. Stein, On the maximal ergodic theorem. Proc. Nat. Acad. Sci. U.S.A. **47** (1961), pp 1894-1897.
[SW78] E. M. Stein, and S. Wainger, Problems in harmonic analysis related to curvature. Bull. Amer. Math. Soc. **84** (1978), pp. 1239–1295.
[T67] A. A. Tempel'man, Ergodic theorems for general dynamical systems. Dokl. Akad. Nauk SSSR 176 (1967), 790793. English translation : Soviet Math. Dokl. 8 (1967), 12131216.
[T92] A. Tempelman, Ergodic theorems for group actions. Mathematics and its applications **78**, Kluwer Academic publishers (1992).
[V] P. Varjú, Random walks in Euclidean space. arXiv:1205.3399.
[vN32] J. von Neumann, Proof of the quasi-ergodic hypothesis, Proc. Nat. Acad. Sci. USA 18 (1932), 70–82.
[W99] M. Waldschmidt, Density measure of rational points on Abelian varieties, Nagoya. Math. J. 155 (1999), 27–53.
[W39] N. Wiener, The ergodic theorem. Duke Math. J. 5 (1939), 1–18.

School of Mathematics, University of Bristol, Bristol, U.K.
E-mail address: `a.gorodnik@bristol.ac.uk`

Department of Mathematics, Technion, Israel
E-mail address: `anevo@tx.technion.ac.il`